

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 16, 2003, 22:42:04 ; Search time 7008.05 Seconds
(without alignments)
17408.415 Million cell updates/sec

Title: US-09-720-086-1
Perfect score: 4192
Sequence: 1 gaattccgcgtcgtgcgcg.....taaaaaaacatacaaatgtc 4192

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Archived: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenDbml:*

- 1: gb_da:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_ov:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vl:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vl:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pln:*
- 35: em_htg_rtd:*
- 36: em_htg_mam:*
- 37: em_htg_vrt:*
- 38: em_sy:*
- 39: em_higo_hum:*
- 40: em_higo_mus:*
- 41: em_higo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4191	100.0	4192	10 AF068625	AF068625 Mus muscu
2	3982.8	95.0	4094	10 BC007466	BC007466 Mus muscu
3	2877.4	68.6	4258	9 AF331856	AF331856 Homo sapi
4	2394.6	57.1	3005	9 AF067972	AF067972 Homo sapi
5	1551.6	37.0	2191	9 AK025230	AK025230 Homo sapi
6	1369.4	32.7	204534	2 AC118195	AC118195 Mus muscu
7	1369.4	32.7	208874	2 AC111092	AC111092 Mus muscu
8	1331.8	31.8	122936	2 AC120824	AC120824 Rattus no
9	1171.2	27.9	176597	2 AC112586	AC112586 Rattus no
10	1164.8	27.8	1758	9 BC018214	BC018214 Homo sapi
11	1055.2	25.2	117264	2 AC103454	AC103454 Rattus no
12	1015.8	24.2	168651	9 AC009474	AC009474 Homo sapi
13	962	22.9	119630	2 AC112040	AC112040 Rattus no
14	796.2	19.0	4135	10 AF068627	AF068627 Mus muscu
15	794.6	19.0	4163	10 AF151974	AF151974 Mus muscu
16	794.6	19.0	4278	10 AF151970	AF151970 Mus muscu
17	789.6	18.8	4267	9 AF176228	AF176228 Homo sapi
18	759.4	18.1	4195	10 AF068626	AF068626 Mus muscu
19	757.8	18.1	4223	10 AF151973	AF151973 Mus muscu
20	757.8	18.1	4338	10 AF151969	AF151969 Mus muscu
21	749.8	17.9	4145	9 AF156488	AF156488 Homo sapi
22	749.8	17.9	4335	9 AF331857	AF331857 Homo sapi
23	640.4	15.3	197652	9 AC012074	AC012074 Homo sapi
24	616.4	14.7	210269	2 AC116459	AC116459 Mus muscu
25	604.8	14.4	3946	10 AF068628	AF068628 Mus muscu
26	603.2	14.4	3974	10 AF151976	AF151976 Mus muscu
27	603.2	14.4	4089	10 AF151971	AF151971 Mus muscu
28	597.6	14.3	3897	9 AF156487	AF156487 Homo sapi
29	568	13.5	4006	10 AY078427	AY078427 Mus muscu
30	566.4	13.5	4034	10 AF151975	AF151975 Mus muscu
31	566.4	13.5	4149	10 AF151972	AF151972 Mus muscu
32	564.6	13.5	119630	2 AC112040	AC112040 Rattus no
33	531.8	12.7	2848	5 AF135438	AF135438 Danio rer
34	482.8	11.5	3017	9 AK001191	AK001191 Homo sapi
35	417.6	10.0	2127	3 AF129267	AF129267 Homo sapi
36	416.4	9.9	167568	2 AC111734	AC111734 Rattus no
37	324.6	7.7	2077	6 AR129189	AR129189 Sequence
38	302.8	7.2	2057	9 AF129268	AF129268 Homo sapi
39	273	6.5	204534	2 AC118195	AC118195 Mus muscu
40	273	6.5	208874	2 AC111092	AC111092 Mus muscu
41	258	6.2	185413	2 AC119009	AC119009 Rattus no
42	255.2	6.1	2008	9 AF129269	AF129269 Homo sapi
43	216.4	5.2	64045	2 AC101453	AC101453 Mus muscu
44	141	3.4	1705	9 AF194032	AF194032 Homo sapi
45	140	3.3	1397	9 BC002560	BC002560 Homo sapi

ALIGNMENTS

RESULT 1
LOCUS AF068625
DEFINITION Mus musculus DNA cytosine-5 methyltransferase 3A (Dnmt3a) mRNA,
complete cds.
ACCESSION AF068625
VERSION AF068625.2
KEYWORDS GI:6449467
SOURCE Mus musculus.
ORGANISM Mus musculus.

REFERENCE 1 (bases 1 to 4192)
Okano, M., Xie, S. and Li, E.
TITLE Cloning and characterization of a family of novel mammalian DNA

```

(cytosine-5) methyltransferases
JOURNAL Nat. Genet. 19 (3), 219-220 (1998)
MEDLINE 98324766
PUBMED 9662389
REFERENCE 2 (bases 1 to 4192)
AUTHORS Xie,S., Okano,M. and Li,E.
TITLE Direct Submision
JOURNAL Submitted (28-MAY-1998) CVRC, Mass. Gen. Hospital, 149 13th Street,
          Charlestown, MA 02129, USA
REFERENCE 3 (bases 1 to 4192)
AUTHORS Okano,M., Chikawa,T., Sasaki,H. and Li,E.
JOURNAL Direct Submision
          Submitted (04-NOV-1999) CVRC, Mass. Gen. Hospital, 149 13th Street,
          Charlestown, MA 02129, USA
REMARK Sequence update by submltler
FEATURES
        source
            1. .4192
                /organism="Mus musculus"
                /db_xref="taxon:10090"
                /chromosome="12"
                /map="4.0 cM"
                1. .4192
                /gene="Dnmt3a"
                1. .4192
                /gene="Dnmt3a"
                /note="similar to EST sequences deposited in GenBank
                Accession Numbers AA052791, AA111043, AA154890, AA240794,
                AA756653, W58898, W59299, W91664, and W91665"
                217. .2943
                /gene="Dnmt3a"
                /function="de novo DNA methylation"
                /note="contains Cys-rich region"
                /product="DNA cytosine-5 methyltransferase 3A"
                /protein_id="AAC40177.2"
                /db_xref="gi:6449468"
                /translation="MPSGPGDPTSSSTIREDRKRGEGROENKKEERDEPESATRK
                VGRPRKRPKHPVSSDPRKDPATVATKESGPMADGSPDLNEDLKRSEPPES
                PAAQKGAPEAGETETPEPAVAENKQCYKESKRAKDELMKREKRAKTYA
                GSRGLRGSLNBSLNQKRPFLITQAGPPTISKRAKDELMKREKRAKTYA
                VMAHVENQASQSKVKEASPRVQDPTDPASPTVATPEVQGDGADNKAADY
                EPEIDEGRGELGELWGLKRGSMWGRIVSMWGRIVSMWGRIVSMWGRIVSMWGRIV
                VCEVKMLPUSFCSAFHQAIVNKPRAIYEVQVASSAGLFPACHDSQSDSG
                KAVEYONQMIEMALGFPQSGRGLPEPEERKNPKYVTDMMVEEAAVAPPPA
                KAPRKSTTEKRAVEIDERTRERLVYEVORCRNIEDICISGLVLTLEHDLFTGG
                MCQCNKCFLECAVQDDGYOSYCTCGGREGVLMQNNCCRCFCVECDLVLPQ
                AAQAIKEDPMNMCYGHKGTGLLRREDWPSRLQMFANNHDEDFDPKVPVPA
                EKRPRIYLSLFDGIATGLYLKLDLQVDRYIASEVEDSITVGMVHQKIMYDQ
                VRSVYQKHIOEMGFEDLYIGSPNDLSIVNPARGLVEGRLFEFFYRLIDAPRK
                EDDDPFVMEVNVAMGVSDKRDISRPLESNPVMIDAEVSAHARARYMGMLPQM
                RPLASTVNDKLELOELEGRIAKFSKVTITTRNSNISKQGDHPFVFNNEEDITM
                CTEMERVGFVHYTVDVSNMSRLARORLGRMSVPIRHLFAFLKRFACV"
BASE COUNT      1094 a      1072 c      1186 g      840 t
ORIGIN
Query Match      100.0%; Score 4191; DB 10; Length 4192;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4191; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

|||||
181 CCTTCACACCCACACCTTGGCCCTACTGCGCAATGCCCTCCAGGGCCCGGGGAC 240
|||||
241 ACCAGAGCTCTCTCTCTGAGAGAGAGAGATGATCAAGAAAGAGAGAGAGAGAG 300
|||||
241 ACCAGAGCTCTCTCTCTGAGAGAGAGAGATGATCAAGAAAGAGAGAGAGAGAG 300
|||||
301 AACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
|||||
301 AACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
|||||
361 GCGCGAAGCGCAAGACACCCACCGGTGAAAAGCATGACACCCCAAGACAGCAGTG 420
|||||
361 GCGCGAAGCGCAAGACACCCACCGGTGAAAAGCATGACACCCCAAGACAGCAGTG 420
|||||
421 ACCACCAAGTCTCAGCCCATGAGCCAGAGACTGTGGCCCTCAGATCTTGACCAATGGA 480
|||||
421 ACCACCAAGTCTCAGCCCATGAGCCAGAGACTGTGGCCCTCAGATCTTGACCAATGGA 480
|||||
481 GACTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
|||||
481 GACTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
|||||
541 GGTGGGGCCCGCGCTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
|||||
541 GGTGGGGCCCGCGCTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
|||||
601 GAGAATGGCTGTGTGTGAGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
|||||
601 GAGAATGGCTGTGTGTGAGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
|||||
661 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
|||||
661 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
|||||
721 GCGTTGGCTGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
|||||
721 GCGTTGGCTGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
|||||
781 GGGAGACCCCTACTACTATCAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
|||||
781 GGGAGACCCCTACTACTATCAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
|||||
841 GAGCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
|||||
841 GAGCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
|||||
901 TCTGAGAGAGTCTCAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
|||||
901 TCTGAGAGAGTCTCAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
|||||
961 CCGTGTCTCCGAGCTGTGGCCACACCCCTGAGAGCAGTAGAGAGAGAGAGAGAG 1020
|||||
961 CCGTGTCTCCGAGCTGTGGCCACACCCCTGAGAGCAGTAGAGAGAGAGAGAGAG 1020
|||||
961 CCGTGTCTCCGAGCTGTGGCCACACCCCTGAGAGCAGTAGAGAGAGAGAGAGAG 1020
|||||
1021 AATGCTACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
|||||
1021 AATGCTACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
|||||
1081 GAGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140
|||||
1081 GAGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140
|||||
1141 TGGTGATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
|||||
1141 TGGTGATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
|||||
1201 GATGGCAAGTCTCAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260
|||||
1201 GATGGCAAGTCTCAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260
|||||
1261 GATTTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
|||||
1261 GATTTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320

```

Db 1261 GCATTCCACGAGGCCACCTACACAGACGCCATGTACCGCAAAAGCCATCTAGCAGATC 1320
Oy 1321 CTCAGGTGGCAGAGCCGTGGCGGGAAGCTGTTTCCAGCTTGCCATGACAGTATGAA 1380
Db 1321 CTCAGGTGGCAGAGCCGTGGCGGGAAGCTGTTTCCAGCTTGCCATGACAGTATGAA 1380
Oy 1381 AGTGACACTGGCAGAGGCTGTGGAAGTGCAGAACAGCAGATGATTTGAATGGGCCCTCGGT 1440
Db 1381 AGTGACACTGGCAGAGGCTGTGGAAGTGCAGAACAGCAGATGATTTGAATGGGCCCTCGGT 1440
Oy 1441 GCGTTCACGCCCTGGGGTCTTAAGGGCTTGCAGCCACAGAGAGAGAGATCTTAC 1500
Db 1441 GCGTTCACGCCCTGGGGTCTTAAGGGCTTGCAGCCACAGAGAGAGATCTTAC 1500
Oy 1501 AAGGAAGTTTACCCGACATGTGGGTGAGGCTGAGAGAGCGCTTACGCCCAACCCCA 1560
Db 1501 AAGGAAGTTTACCCGACATGTGGGTGAGGCTGAGAGAGCGCTTACGCCCAACCCCA 1560
Oy 1561 CCAGCCAGAAACCCAGAAAGAGCACAACAGAGAACTTAAGTCAAGAGATCAATTGAT 1620
Db 1561 CCAGCCAGAAACCCAGAAAGAGCACAACAGAGAACTTAAGTCAAGAGATCAATTGAT 1620
Oy 1621 GAGCGCACAGAGAGCGCGTGTATGAGTGGCCAGAAAGTGCAGAAACATCGAGAC 1680
Db 1621 GAGCGCACAGAGAGCGCGTGTATGAGTGGCCAGAAAGTGCAGAAACATCGAGAC 1680
Oy 1681 ATTGTATCTATGTGGAGGCTCAATGTCAACCCCTGGAGCACCCCTCTTCAATGAGAGC 1740
Db 1681 ATTGTATCTATGTGGAGGCTCAATGTCAACCCCTGGAGCACCCCTCTTCAATGAGAGC 1740
Oy 1741 ATGTGCCAGAACTGTAAAGTGTCTTGTGAGTGTGCTTACCAATATGAGCAGATGGG 1800
Db 1741 ATGTGCCAGAACTGTAAAGTGTCTTGTGAGTGTGCTTACCAATATGAGCAGATGGG 1800
Oy 1801 TACCACTCTTATGACCATCTGCTGTGGGGGGGTGAAGTGTCTATGTGTGGAACAC 1860
Db 1801 TACCACTCTTATGACCATCTGCTGTGGGGGGGTGAAGTGTCTATGTGTGGAACAC 1860
Oy 1861 AACCTCTGAGTGTGTGTGTGAGTGTGTGATCTTGTGTGGGGGCGAGAGCTGCT 1920
Db 1861 AACCTCTGAGTGTGTGTGTGAGTGTGTGATCTTGTGTGGGGGCGAGAGCTGCT 1920
Oy 1921 CAGCGAGCCATTAAAGAAAGACCCCTGGAATGCTACATGTGGGGCATTAAGGGCACCTAT 1980
Db 1921 CAGCGAGCCATTAAAGAAAGACCCCTGGAATGCTACATGTGGGGCATTAAGGGCACCTAT 1980
Oy 1981 GGGCTGTGCGGAAGACGGGAAGACTGGCTTCTGCACATCCAGATTTCTTTGCCAATAAC 2040
Db 1981 GGGCTGTGCGGAAGACGGGAAGACTGGCTTCTGCACATCCAGATTTCTTTGCCAATAAC 2040
Oy 2041 CATGACCAAGAAATTTGACCCCAAAAGTTTACCACTGTGCCAGCTGAGAGAGAGAG 2100
Db 2041 CATGACCAAGAAATTTGACCCCAAAAGTTTACCACTGTGCCAGCTGAGAGAGAGAG 2100
Oy 2101 CCCATCCGCTGTCTCTCTTTGTGATGGATTTGTACAGAGGCTCTGGTGTGAAGAC 2160
Db 2101 CCCATCCGCTGTCTCTCTTTGTGATGGATTTGTACAGAGGCTCTGGTGTGAAGAC 2160
Oy 2161 CTGGCATTCAGAGTGGACCGCTACATTTGCTTCCAGAGTGTGTGAGAGACTCCATCAGCGTG 2220
Db 2161 CTGGCATTCAGAGTGGACCGCTACATTTGCTTCCAGAGTGTGTGAGAGACTCCATCAGCGTG 2220
Oy 2221 GGCATGTGGGGCACAGGGAAGATCATGTACGTGGGGAGAGTCCGGAAGGCTCACAG 2280
Db 2221 GGCATGTGGGGCACAGGGAAGATCATGTACGTGGGGAGAGTCCGGAAGGCTCACAG 2280
Oy 2281 AAGCATATCCAGAGTGGGGCCATTCGACCTGTGTATGAGAGAGTCCCTGCAATGAC 2340
Db 2281 AAGCATATCCAGAGTGGGGCCATTCGACCTGTGTATGAGAGAGTCCCTGCAATGAC 2340
Oy 2341 CTCCTCATTTGTAACCCCTGCGGCAAGGACTTTATGAGAGTACTGGCCCTCTTCTTT 2400
Db 2341 CTCCTCATTTGTAACCCCTGCGGCAAGGACTTTATGAGAGTACTGGCCCTCTTCTTT 2400

Oy 2401 GAGTTCATCCGCTCTGATGATGCGGGCCAAAGAGAGATATCGCCCTTCTTC 2460
Db 2401 GAGTTCATCCGCTCTGATGATGCGGGCCAAAGAGAGATATCGCCCTTCTTC 2460
Oy 2461 TGGCTCTTTGAGAAATGTGGTGGCCATTTGGGCTTGTGACAAAGAGGACATCTCGGATTT 2520
Db 2461 TGGCTCTTTGAGAAATGTGGTGGCCATTTGGGCTTGTGACAAAGAGGACATCTCGGATTT 2520
Oy 2521 CTGAGTCTAACCCCGTATGATTCAGCCCAAGAGTGTGCTGACACAGGGCCGT 2580
Db 2521 CTGAGTCTAACCCCGTATGATTCAGCCCAAGAGTGTGCTGACACAGGGCCGT 2580
Oy 2581 TACTTCTGGGTAACTCTCTGCGATGAAACAGGCTTTGGCATCCATGTAATGATAG 2640
Db 2581 TACTTCTGGGTAACTCTCTGCGATGAAACAGGCTTTGGCATCCATGTAATGATAG 2640
Oy 2641 CTGAGCTGCAAGAGTGTGTGAGCAGCGGCAATATGACCAAGTTGACCAAGTGAAGAC 2700
Db 2641 CTGAGCTGCAAGAGTGTGTGAGCAGCGGCAATATGACCAAGTTGACCAAGTGAAGAC 2700
Oy 2701 ATTACCACTGATCAACTCTTATTAAGCAGAGGCAAGACAGCATTTCCCGCTTCAATG 2760
Db 2701 ATTACCACTGATCAACTCTTATTAAGCAGAGGCAAGACAGCATTTCCCGCTTCAATG 2760
Oy 2761 AACGAGAGAGACATCTCTGTGTGCACTGAAATGAAAGGGTGTGGCTTCCCGCTC 2820
Db 2761 AACGAGAGAGACATCTCTGTGTGCACTGAAATGAAAGGGTGTGGCTTCCCGCTC 2820
Oy 2821 CACTACAGAGAGCTGCAACATGAGCCGCTTGGAGGACAGACCTGTGGCCGATCG 2880
Db 2821 CACTACAGAGAGCTGCAACATGAGCCGCTTGGAGGACAGACCTGTGGCCGATCG 2880
Oy 2881 TGGAGCTGCGGTATCCGCCACTCTTCCGTCCGCTGGAAGATATTTTCTGTGTG 2940
Db 2881 TGGAGCTGCGGTATCCGCCACTCTTCCGTCCGCTGGAAGATATTTTCTGTGTG 2940
Oy 2941 TAAAGACATGGGGCAAACTGAATGATGATGATGATGATGATGATGATGATGATGATG 3000
Db 2941 TAAAGACATGGGGCAAACTGAATGATGATGATGATGATGATGATGATGATGATGATG 3000
Oy 3001 AAAAAACAACAAATTAACACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060
Db 3001 AAAAAACAACAAATTAACACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060
Oy 3061 CAG 3120
Db 3061 CAG 3120
Oy 3121 CTTCGAAAAAGGTTGAGACATCTCTGAGTGTGCAATGTTAACCTTCACTATCTA 3180
Db 3121 CTTCGAAAAAGGTTGAGACATCTCTGAGTGTGCAATGTTAACCTTCACTATCTA 3180
Oy 3181 AAAAGCAAAATATGAGCCCTCCCTCTTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 3240
Db 3181 AAAAGCAAAATATGAGCCCTCCCTCTTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 3240
Oy 3241 TTCTACTCTTTTTCAGAGGGGTTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3300
Db 3241 TTCTACTCTTTTTCAGAGGGGTTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3300
Oy 3301 AAAAGAGAGAGTTATTCAGCAAAATAGTAACAACAAAGTAAAGTAAAGTAAAGTAAAG 3360
Db 3301 AAAAGAGAGAGTTATTCAGCAAAATAGTAACAACAAAGTAAAGTAAAGTAAAGTAAAG 3360
Oy 3361 GAAAGGAG 3420
Db 3361 GAAAGGAG 3420
Oy 3421 TTTCTATATATCTCTTGT 3480
Db 3421 TTTCTATATATCTCTTGT 3480

OY	4481	AATAAGAGCCCTGGAGGGAGAGTCTCTCCACCCCGAGAGCTCAACAGCAC	3548
Db	3481	AATAAGAGCCCTGGAGGGAGAGTCTCTCCACCCCGAGAGCTCAACAGCAC	3540
OY	3541	ATTCTGGTATFCAAAACAGAAACCAACTAGCAGAGGGCGGTGAGAGAAACACAC	3600
Db	3541	ATTCTGGTATFCAAAACAGAAACCAACTAGCAGAGGGCGGTGAGAGAAACACAC	3600
OY	3601	AGACACTTTCACAGATATTTGAGGGCCCTGACACAGAGAAACCTGAGAAAACCAAGT	3660
Db	3601	AGACACTTTCACAGATATTTGAGGGCCCTGACACAGAGAAACCTGAGAAAACCAAGT	3660
OY	3661	TCTAGAAAGCCGCTTACTCTGTGTTACAGTTTATATATATATAGATAGATATAT	3720
Db	3661	TCTAGAAAGCCGCTTACTCTGTGTTACAGTTTATATATATATAGATAGATATAT	3720
OY	3721	ATAATATATAAAGTACGTGTTAACACTGACATCCCGACTCAATAATGGTGTTCAAA	3780
Db	3721	ATAATATATAAAGTACGTGTTAACACTGACATCCCGACTCAATAATGGTGTTCAAA	3780
OY	3781	ACACGAGATAGACAAACAGACATCAGCTTCCCGCTGGCCCTGTGTGCAAGGGTTTCAGCC	3840
Db	3781	ACACGAGATAGACAAACAGACATCAGCTTCCCGCTGGCCCTGTGTGCAAGGGTTTCAGCC	3840
OY	3841	CAGATATGGGAGAGGGGGAGCAGCTGTGAGGGGGTTTTAACAAATGAAAGATGACCCATAT	3900
Db	3841	CAGATATGGGAGAGGGGGAGCAGCTGTGAGGGGGTTTTAACAAATGAAAGATGACCCATAT	3900
OY	3901	CACCCCCCACTCCCTGCCCATGCTCCTAGCTTCACTCCGCAAAAGAGGGCTCAGCTGAGGTG	3960
Db	3901	CACCCCCCACTCCCTGCCCATGCTCCTAGCTTCACTCCGCAAAAGAGGGCTCAGCTGAGGTG	3960
OY	3961	GTGCGAGCCCTGGGGAAAGCTGAGTGTGGAATTTATCCAGACTCGCGTCAATAACTTGA	4020
Db	3961	GTGCGAGCCCTGGGGAAAGCTGAGTGTGGAATTTATCCAGACTCGCGTCAATAACTTGA	4020
OY	4021	ATATGATATCTAAATATGACTGCCCTAGAAATAATGGCTTGAGAAAACATGTCCCTGATTTT	4080
Db	4021	ATATGATATCTAAATATGACTGCCCTAGAAATAATGGCTTGAGAAAACATGTCCCTGATTTT	4080
OY	4081	GAATTCGTGACGCCAGCTTGAAGGCCCTTGTGGGATCGAAATATTCACAGGTGAGGAA	4140
Db	4081	GAATTCGTGACGCCAGCTTGAAGGCCCTTGTGGGATCGAAATATTCACAGGTGAGGAA	4140
OY	4141	AGTAGCCGCGCATTAACCCGCTGGAGCAATATATAAATACATCAAAATGT	4192
Db	4141	AGTAGCCGCGCATTAACCCGCTGGAGCAATATATAAATACATCAAAATGT	4192

RESULT 2	BC007466	4094 bp	mRNA	linear	ROD 07-AUG-2002
DESCRIPTION	BC007466	Mus musculus, DNA methyltransferase 3A, clone MGC:55662			
ACCESSION	BC007466	IMAGE:3492853, mRNA, complete cds.			
VERSION	BC007466.1	GI:13938620			
KEYWORDS	MGC.				
SOURCE	house mouse.				
ORGANISM	Mus musculus.				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	1. (bases 1 to 4094)				
JOURNAL	Strausberg, R.				
TITLE	Direct Submision				
REMARK	Submitted (01-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				
COMMENT	NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: cgapsb-remail.nih.gov Tissue Procurement: Gilbert Smith, Ph.D. cDNA Library Preparation: Life Technologies, Inc.				

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>
Contact: (Dickson, Mark) mdc@paxll.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MCC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Series: IRK Plate: 8 Row: n Column: 11
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6661208.
Location/Qualifiers
1..4094
source

CDS

1067 a 1057 c 1164 g 806 t

Query Match	95.0%	Score 3982.8	DB 10	Length 4094	
Best Local Similarity	99.6%	Pred. No. 0			
Matches 4035	Conservative	0	Mismatches 12	Indels 5	Gaps 4
QY	16	GCCGGGCCCCGACCGGGGCGACACAGGAGAGCCGGCTGAAACCCAGGCGTAGGC	75		
Db	44	GCCGGGCGCCGACCGCGGGGCGACAGGGAGAGCCGGCTGAAACCCAGGCGTAGGC	103		
QY	76	TGACATTTTCGAGGGGCTTGACATCAAGGGTCATGTTAAGTCCTTGAAGCTCTTGA	155		
Db	104	TGACATTTTCGAGGGGCTTGACATCAAGGGTCATGTTAAGTCCTTGAAGCTCTTGA	163		
QY	136	AGACACAGGGCAATTCCTTCCTGTAACCCCTCGACACCCGCACAGCCGCTCGAGCCCG	195		
Db	164	AGACACAGGGCAATTCCTTCCTGTAACCCCTCGACACCCGCACAGCCGCTCGAGCCCG	223		
QY	196	CTGCGCGCTTACTGCGCAGCAATGCCCCCTCGACGGCCCCGGGACACACAGCACTCTCT	255		
Db	224	CTGCGCGCTTACTGCGCAGCAATGCCCCCTCGACGGCCCCGGGACACACAGCACTCTCT	283		
QY	256	CTGAGCGCGAGAGATGATCGAAAGAGAGAGACAGAGAGAGAACTGTGCAAGGAA	315		
Db	284	CTGAGCGCGAGAGATGATCGAAAGAGAGAGACAGAGAGAGAACTGTGCAAGGAA	343		

316 GAGCCGACAGAGCCCGAGCCGACGGCCGGAAGGTTGGGAGGCTTGGCCGGAAGGCGAAG 375
|||||
344 GAGCCGACAGAGCCCGAGCCGACGGCCCGGAAGGTTGGGAGGCTTGGCCGGAAGGCGAAG 403
|||||
376 CACCCACCGGTTGGAAGAGTACACCCGACAGGAGCCAGTACAGCACCAGTCTGAG 435
|||||
404 CACCCACCGGTTGGAAGAGTACACCCGACAGGAGCCAGTACAGCACCAGTCTGAG 463
|||||
436 CCCATGGCCGACAGTCTGGCCCTCAGATCTGCTACCCCATGAGACTTGGAGAAAGCGG 495
|||||
464 CCCATGGCCGACAGTCTGGCCCTCAGATCTGCTACCCCATGAGACTTGGAGAAAGCGG 523
|||||
496 AGTGAACCCCAACCTGAGAGGAGGAGCCAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 555
|||||
524 AGTGAACCCCAACCTGAGAGGAGGAGGAGCCAGCTGAGAGGAGGAGGAGGAGGAGGAGGAG 583
|||||
556 GAAGCAGAGGAGAACTGAGACCCGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 615
|||||
584 GAAGCAGAGGAGAACTGAGACCCGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 643
|||||
616 GTGACCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 675
|||||
644 GTGACCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 703
|||||
676 ATCGAATCCATGAAATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 735
|||||
704 ATCGAATCCATGAAATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 763
|||||
736 TCACAGCTCTGCTCAGAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 795
|||||
764 TCACAGCTCTGCTCAGAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 823
|||||
796 ATCGCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 855
|||||
824 ATCGCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 883
|||||
856 GCCAAGGTAATTTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 915
|||||
884 GCCAAGGTAATTTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 943
|||||
916 AAGGTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 975
|||||
944 AAGGTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1003
|||||
976 GTGGCCACACCCCTGAGCAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1035
|||||
1004 GTGGCCACACCCCTGAGCAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1063
|||||
1036 GCCGACGATGAGCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1095
|||||
1064 GCCGACGATGAGCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1123
|||||
1096 GGGAAACTTGGGGGCTTCTCTGGTGGCCAGGCGCAATTTGCTTGGTGGATGACAGGCG 1155
|||||
1124 GGGAAACTTGGGGGCTTCTCTGGTGGCCAGGCGCAATTTGCTTGGTGGATGACAGGCG 1183
|||||
1156 CGGAGCCGAGCAGCTGAGAGGAGCAGTGGTGGATGATGATGATGATGATGATGATGATGATGAT 1215
|||||
1184 CGGAGCCGAGCAGCTGAGAGGAGCAGTGGTGGATGATGATGATGATGATGATGATGATGATGAT 1243
|||||
1216 GTGGTGTGTGGAGAACTATGCGCGCTGAGCCTCTCTGCACTGCAATTCACACAGGCG 1275
|||||
1244 GTGGTGTGTGGAGAACTATGCGCGCTGAGCCTCTCTGCACTGCAATTCACACAGGCG 1303
|||||
1276 ACCTTACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1335
|||||
1304 ACCTTACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1363
|||||
1336 AGCCCTGCGGGAAGCTTTTCCAGCTTGGCAGTACAGTGAAGTGAAGTGAAGTGAAGTGAAG 1395
|||||
1364 AGCCCTGCGGGAAGCTTTTCCAGCTTGGCAGTACAGTGAAGTGAAGTGAAGTGAAGTGAAG 1423
|||||

QY 1396 GCTGTGGAAGTGCACAAACAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1455
|||||
Db 1424 GCTGTGGAAGTGCACAAACAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1483
|||||
QY 1456 GGTCTTAAGGAGCCTGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1515
|||||
Db 1484 GGTCTTAAGGAGCCTGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1543
|||||
QY 1516 GACATGTGGTGGAGCCTGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1575
|||||
Db 1544 GACATGTGGTGGAGCCTGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1603
|||||
QY 1576 AGAAAGACACACAGAGAACTTAAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1635
|||||
Db 1604 AGAAAGACACACAGAGAACTTAAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1663
|||||
QY 1636 CGGCTGTGTATGAGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1695
|||||
Db 1664 CGGCTGTGTATGAGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1723
|||||
QY 1696 GGGAGCCTCAATGTCAACCTGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1755
|||||
Db 1724 GGGAGCCTCAATGTCAACCTGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1783
|||||
QY 1756 AAGAACTGCTTCTTGGAGTGTGCTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1815
|||||
Db 1784 AAGAACTGCTTCTTGGAGTGTGCTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1843
|||||
QY 1816 ACCATCTGCTGTGGGGGAGGTGAAGTGTCTATGTGTGGGAACACACACTGTCTGAGGTGC 1875
|||||
Db 1844 ACCATCTGCTGTGGGGGAGGTGAAGTGTCTATGTGTGGGAACACACACTGTCTGAGGTGC 1903
|||||
QY 1876 TTTTGTGTGAGATGT 1935
|||||
Db 1904 TTTTGTGTGAGATGT 1963
|||||
QY 1936 GAAAGACCCCTGAACTGTCTACATGT 1995
|||||
Db 1964 GAAAGACCCCTGAACTGTCTACATGT 2023
|||||
QY 1996 CGGGAAGACTGGCCTTCTGAGCTCAGATGTTCTTGGCAATTAACCATGACAGGAATTT 2055
|||||
Db 2024 CGGGAAGACTGGCCTTCTGAGCTCAGATGTTCTTGGCAATTAACCATGACAGGAATTT 2083
|||||
QY 2056 GACCCGCCAAAGGTTTACCCACCTGTCAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2115
|||||
Db 2084 GACCCGCCAAAGGTTTACCCACCTGTCAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2143
|||||
QY 2116 TCTCTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2175
|||||
Db 2144 TCTCTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2203
|||||
QY 2176 GACCGCTACATTTGCTCCGAGGTGTGTGAGAGCTCCATTCACGCTGGGATGAGTGGCGAC 2235
|||||
Db 2204 GACCGCTACATTTGCTCCGAGGTGTGTGAGAGCTCCATTCACGCTGGGATGAGTGGCGAC 2263
|||||
QY 2236 CAGGAAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2295
|||||
Db 2264 CAGGAAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2323
|||||
QY 2296 TGGGGCCATTCGACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2355
|||||
Db 2324 TGGGGCCATTCGACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2383
|||||
QY 2356 CCTGCGCGCAGAGGACTTTATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2415
|||||
Db 2384 CCTGCGCGCAGAGGACTTTATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2443
|||||
QY 2416 CTGCATGATGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2475
|||||
Db 2444 CTGCATGATGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2503
|||||
QY 2476 GTGGTGGCCATGAGGCGTTAGTGAACAAAGAGGAGCATTCTGCGAATTTCTTGAGTCTAACC 2535
|||||

Db	2504	GTGATGTCATGAGGCGCTTATGTCACAGAGGAGCATCTGCGATTTCTGATGCTTAACCC	2563	Db	3580	AAACAGAACCCAACTAGCAGCAGGCGCTGAGAGAACACCACACGACACTTTTCTACA	3639
QY	2536	GTGATGATGAGCCCAAGAAAGTGTCTCTGACACAGGCGCCCTTAATCTTGGGGTAAC	2595	QY	3615	GTATTTAGGTGCTTACACACAGAGAAACCTTGAGAGAAACAGCTTTCTAGAACCGCTG	3674
Db	2564	GTGATGATGAGCCCAAGAAAGTGTCTCTGACACAGGCGCCCTTAATCTTGGGGTAAC	2623	Db	3640	GTATTTAGGTGCTTACACACAGAGAAACCTTGAGAGAAACAGCTTTCTAGAACCGCTG	3699
QY	2596	CTTCTGTCATGAGCAGGCGCTTTGGCATCTGATGATGATGATGATGATGATGATGATG	2655	QY	3675	TTAAGCTTCTGTTTACAGTTTATATATATATATATATATATATATATATATATATAT	3734
Db	2624	CTTCTGTCATGAGCAGGCGCTTTGGCATCTGATGATGATGATGATGATGATGATGATG	2683	Db	3700	TTAAGCTTCTGTTTACAGTTTATATATATATATATATATATATATATATATATATAT	3759
QY	2656	TGCTGAGAGCAGCAGAAATAGCAGAAAGTTCAGCAAAAGTTCAGCAAAAGTTCAGCA	2715	QY	3735	TAGCTTAAGCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	3794
Db	2684	TGCTGAGAGCAGCAGAAATAGCAGAAAGTTCAGCAAAAGTTCAGCAAAAGTTCAGCA	2743	Db	3760	TAGCTTAAGCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	3819
QY	2716	AACTCTATTAAGCAGGCGCAAGACGATTTCCCGCTTTTCATGAAAGAGAGAGAGAG	2775	QY	3795	AAAGACATCAGCTTCCCGCTGCGCTGCTGTCGCAAAAGGCTTTCAGCCAGATGGGAG	3854
Db	2744	AACTCTATTAAGCAGGCGCAAGACGATTTCCCGCTTTTCATGAAAGAGAGAGAGAG	2803	Db	3820	AAAGACATCAGCTTCCCGCTGCGCTGCTGTCGCAAAAGGCTTTCAGCCAGATGGGAG	3879
QY	2776	ATCTGTCGTCATGAAATGGAAGGCTTTGGCTTCCCGCTTCACAGCAGCAGAGCTC	2835	QY	3855	GAGAGCAGCTGAGAGGCGGTTTAAACAGAGATGAGCAGATATACACCCACCCCT	3914
Db	2804	ATCTGTCGTCATGAAATGGAAGGCTTTGGCTTCCCGCTTCACAGCAGCAGAGCTC	2863	Db	3880	GAGAGCAGCTGAGAGGCGGTTTAAACAGAGATGAGCAGATATACACCCACCCCT	3939
QY	2836	TCACACATGAGCGCTTGGCGAGGACAGACTGCTGGGCGGATGCTGAGAGCTGCCGCTC	2895	QY	3915	GCCCCATGCTTACCTTACCTGACCTGCAAAAGGCGCTCAGCTGAGGTGCTGGACCTGGG	3974
Db	2864	TCACACATGAGCGCTTGGCGAGGACAGACTGCTGGGCGGATGCTGAGAGCTGCCGCTC	2923	Db	3940	GCCCCATGCTTACCTTACCTGACCTGCAAAAGGCGCTCAGCTGAGGTGCTGGACCTGGG	3999
QY	2896	ATCCGCAACCTTTCGCTCGCTGAGAGATATTTTCTGCTGCTGCTGCTGCTGCTGCTG	2955	QY	3975	AAAGTGTGTGGAATTTATATGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	4034
Db	2924	ATCCGCAACCTTTCGCTCGCTGAGAGATATTTTCTGCTGCTGCTGCTGCTGCTGCTG	2983	Db	4000	AAAGTGTGTGGAATTTATATGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	4059
QY	2956	CAAACTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	3015	QY	4035	TGACTGCTCTGAGAAATGAGCTTGGAGAAACA	4066
Db	2984	CAAACTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	3043	Db	4060	TGACTGCTCTGAGAAATGAGCTTGGAGAAACA	4091
QY	3016	ACATATTAAGCAGGCGCAAGACGATTTCCCGCTTTTCATGAAAGAGAGAGAGAG	3075	RESULT 3			
Db	3044	ACATATTAAGCAGGCGCAAGACGATTTCCCGCTTTTCATGAAAGAGAGAGAGAG	3103	AF331856			
QY	3076	AATTTAAGCAGCAG	3125	LOCUS	AF331856		
Db	3104	AATTTAAGCAGCAG	3153	DEFINITION	4258 bp mRNA linear PRI 02-JAN-2002		
QY	3136	GACATCATCTCTGATTTTCAATGATGATGATGATGATGATGATGATGATGATGATG	3195	Homo sapiens DNA cytosine methyltransferase 3 alpha (DNMT3A) mRNA,			
Db	3164	GACATCATCTCTGATTTTCAATGATGATGATGATGATGATGATGATGATGATGATG	3223	complete cds.			
QY	3196	CCCTCCCTCTTCCCTCGGCTCTAGAGAGCACTTTTGTGTTTCTTCTCTTCTTCTCA	3255	AF331856			
Db	3224	CCCTCCCTCTTCCCTCGGCTCTAGAGAGCACTTTTGTGTTTCTTCTCTTCTTCTCA	3283	VERSION	AF331856.1 GI:18033252		
QY	3256	GAGGCGTCTTCTGTTTGGGTTTGTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3315	KEYWORDS			
Db	3284	GAGGCGTCTTCTGTTTGGGTTTGTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3343	ORGANISM	Homo sapiens.		
QY	3316	TGCAGCAAAATCAGTAACAACAAAAAGTGAATGCTTGGAGAGAGAGAGAGAGAG	3375	REFERENCE	Homo sapiens.		
Db	3344	TGCAGCAAAATCAGTAACAACAAAAAGTGAATGCTTGGAGAGAGAGAGAGAGAG	3401	AUTHORS	N.J., Pradhan, S. and Roberts, R.J.		
QY	3376	AAATTTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	3435	TITLE	Cloning, expression and characterization of human DNMT3 genes		
Db	3402	AAATTTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	3460	JOURNAL	Unpublished		
QY	3436	TTGGTGTCTCTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG	3495	AUTHORS	N.J., Pradhan, S. and Roberts, R.J.		
Db	3461	TTGGTGTCTCTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG	3520	TITLE	Direct Submission		
QY	3496	AGGAGAGTCTCTCTCCAGCCCGGAGAGTCTCAACAGCAGCCTTCCTGCTATGCA	3555	JOURNAL	Submitted (22-DEC-2000) New England Biolabs, 32 Tozer Road,		
Db	3521	AGGAGAGTCTCTCTCCAGCCCGGAGAGTCTCAACAGCAGCCTTCCTGCTATGCA	3579	FEATURES	Beverly, MA 01915, USA		
QY	3556	AAACAGAACCCAACTAGCAGCAGGCGCTGAGAGAACACCACACGACACTTTTCTACA	3614	source	Location/Qualifiers		
				gene	1. 4258		
				CDS	/organism="Homo sapiens"		
					/db_xref="taxon:9606"		
					1. 4258		
					/gene="DNMT3A"		
					237. 2856		
					/gene="DNMT3A"		
					/codon_start=1		
					/product="DNA cytosine methyltransferase 3 alpha"		
					/protein_id="AA157039.1"		
					/db_xref="GI:18033253"		
					/translation="MPSSGPGDSSSAAREERDKGEDEPRKKEERDEPSTTARK		
					VRRPGRKRHPVSGDTPDPDAVIVSKSPMDGASSELLPNQDLKRSRPPGSGS		
					PAGGQKGAPAGEGGAAETLEASRAVENCCCTPREGAGAEAGKQDKSTNIESKIM		
					ESRRLRGLGWSSSLRRPMPRLTFQAGDPIYISKRRKRWELRWKREARRKARYI		

BASE COUNT	1062 a	1089 c	1257 g	850 t
ORIGIN				

Query Match	68.6%	Score 2877.4	DB 9	Length 4258
Best Local Similarity	84.3%	Pred. No. 0		
Matches 3543	Conservative	0	Mismatches 562	Indels 99
				Gaps 23

Db	74	CAGAGGACCGCTTACCAGAGGGCTCAACACCGGGATTTATGTTTAAGTTTAAGTCTG	133
QY	129	CTTCAAAAGACACGAGCAATCTCTTCTGTGAAGCCCTGCGAGCCCAAGCCCTCGCA	188
Db	134	CTCTCAAAAGACACGATTAATTCCTTCCCAAGGCC-AGCAGCCCCCAGC-CCGCGCCA	191
QY	189	GGCCAGCCTCCGCTACTGSCCA-----GCATTCCTCCAGCGGGCCCCGGGACA	241
Db	192	GGCCAGCCTGCCCTCCGGGCGCCAGATAGCCCGGCATCCTTCCAGCGGGCCCCGGGACA	251
QY	242	CCAGCAGCTCTCTGTGAGCGGGAGATGATCGAAAGAGAGAGAGACAGAGAGAGA	301
Db	252	CCACACACTCTGCTGCGGAGCGGGAGGAGACCAAAAGACGAGAGAGCAGAGAGAGC	311
QY	302	ACGCTGGCAAGAGAGAGGCGCAGGACCCACGCGCAGCGCCCCGGAAAGGTGGGGAGCCTG	361
Db	312	CGCTGGCAAGAGAGAGGCGCAAGGCCACACCGCACCGGAAAGGTGGGGCGCTG	371
QY	362	GCCGAGCGCAACACGACCCACCGGTGGAACATGTACACCCCAAGACCCAGCAGTGA	421
Db	372	GGAGCAAGCGCAACACCCCGGTGGAAACGCGTGAACCGCCAAAGAACCTCTGGGTGA	431
QY	422	CCACCAAGTCTCAGCCCATGGGCCAGGACTGGGCCCTCAGATCTGTAACCAATGGAG	481
Db	432	TCTCCAGTCCCCATTCATGTGCCCCAGGACTCAGCGCCTCAGAGCTATTACCAATGGGG	491
QY	482	ACTTGGAGAGCGGAGTGAACCCCACTGAGGAGGGGAGGCCAGCTCAGAGCGAAGG	541
Db	492	ACTTGGAGAGCGGAGTGAAGCCCCAGCCAGGAGGGGAGGCCCTGTGGGGCGCAAGG	551
QY	542	GTGGGGCCCCAGCTGAAGAGGAGG--ACTGTGAGACCCCAACGAAACCTCTCAGAGCTG	598
Db	552	GGGGGGCCCCAGCAGAGGAGAGGGGTGAGACTTGTCTGAAAGCTCTCAAGACGAC	611
QY	599	TGGGAATGCGTCTGTGTGACCAAGAAAGCCGTGGAGCCTCTCAGAGAGAGGGCAAG	658
Db	612	TGGAATAATGCGTGTGACACCCCAAGGAGGGCCGAGGCCCTGTGAGAAAGGGCAAG	671
QY	659	AACGAAGACGACCAACATTCGATTCGAAATGGAGGGTCCCGGGGGCCGATCGCAG	718
Db	672	AACGAAGAGCAACCAACATTCGAAATTCGAAATGGAGGGTCCCGGGGGCCGCTCGCGG	731
QY	719	GTGACTTGGGTGGGAGTCAGCCTCCGTCAGCGACCCATCCAAAGATCTACCTTCAGG	778
Db	732	GTGCTTGGGTGGGAGTCCAGCCTCCGTCAAGCCGCGCCATCCGAGGCTCACTTCCAG	791
QY	779	CAGGGGACCCCTACTACATCAGCAAGGAAACGGGATGAGTGGCTGGCACGTTGGAAA	838
Db	792	CGGGGGACCCCTACTACATCAGCAAGGCGCAAGCGGAGAGTGGCTGGCACGCTGGAAA	851
QY	839	GGGAGGCTGAGAAAGAAAGCCAAAGTAAATGCAAGTAATGCTGTGGAAAGAACACAGG	898

Db	852	GGAGGGCTGAGAAAGAAAGCCAAAGGTCANTGAGAGATGATATGCTGTGGAAGAAAACCAAG	911
OY	899	CCCTGTGAGAGTCTCAGAGGTGGAGAGGCGACGCCCTCTGCTGTGCAGCAGCCACGG	958
Db	912	GGCCCCGGGAGTCTCAGAAAGGTGGAGAGGCGACGCCCTCTGCTGTGCAGAGGCCACTG	971
OY	959	ACCCCTGCTCTCCCACTGTGTGGCCACCAACCCCTGAGGCCAGTATGAGAGGGATCTCTGGGACA	1018
Db	972	ACCCCGCATCCCACTGTGTGTACCAACGCTTGAGCCCGGTGGGTCTCCATCTCTGGGACA	1031
OY	1019	AGATGTCTACCAAGCAGCCAGCATGAGCTCTGATTTGAGATGAGCGGGGGCTTTTGACA	1078
Db	1032	AGATGTCCACCAAGCAGCGCATATACAGCCAGAGTACAGAGACGGCGGGGGCTTTTGACA	1091
OY	1079	TTGAGAGAGCTGTGTGGGGGAAACTTGGGGGCTTCTCCTGTGTGCACAGGCCGAATGTGT	1138
Db	1092	TTGGGGAGAGCTGTGTGGGGGAAACTTGGGGGCTTCTCCTGTGTGCACAGGCCGCAATGTGT	1151
OY	1139	CTTGTGTGATATACAGGCGCGGAGAGCTGTAAGGACATCTGCTGGGTCAATGTGTTCG	1198
Db	1152	CTTGTGTGATATACGCGCGCGGAGAGAGCTGTAAGGACACCGGTGGGTCAATGTGTTCG	1211
OY	1199	GAGATGGCAATTTCTCAGT	1258
Db	1212	GAGAGGGCAATTTCTCAGT	1271
OY	1259	GTGCATTTCCACGAGGCCACCTTACACAACAGCAGCCCATGTATACCGCAAAAGCCATTCAGAG	1318
Db	1272	GTGCATTTCCACGAGGCCACCTTACACAACAGCAGCCCATGTATACCGCAAAAGCCATTCAGAG	1331
OY	1319	TTCCTCCAGGTGGCCGAGAGCGCGTGGCGGGAAGCTTTTCCAGCTTTGCCATACAGTATG	1378
Db	1332	TTCCTCCAGGTGGCCGAGAGCGCGTGGCGGGAAGCTTTTCCAGCTTTGCCATACAGTATG	1391
OY	1379	AAAGTGCAGTGGCAAGGCTGTGGAAATGTGCAGAAACAGCAATGATTAATGGGAGCCCTCG	1438
Db	1392	AGAGTGCACATCTCCCAAGGCCGCTGGAGGTGTGCAGAAACAGCCCATGATTAATGGGAGCCCTCG	1451
OY	1439	GTGCCTTCACAGCCCTCGGGGCTCTTAAGGGGCTGTGAGGCCACCGAAGAAAGAAATCCTT	1498
Db	1452	GGGGCTTTCACAGCCCTTCCTGGCCCTTAAGGGGCTGTGAGGCCACCGAAGAAAGAAATCCTT	1511
OY	1499	ACAAGGAAGTTTACACGCATATGT	1558
Db	1512	ACAAGGAAGTTTACACGCATATGT	1571
OY	1559	CACCAAGCAAAAACCCAGAAAGAGCAACAAAGAAACCTTAAGTCAAGAGATCATTTG	1618
Db	1572	CACCAAGCAAAAACCCCGGAGAGACACAGCGGAGAAAGCCCAAGGTCAAGAGATTTATTTG	1631
OY	1619	ATGAGCGCACAAAGGGAGCGGCTGTGTATGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1678
Db	1632	ATGAGCGCACAAAGGGAGCGGCTGTGTATGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1691
OY	1679	ACAATTTGATCTCATGT	1738
Db	1692	ACAATCTGATCTCTGT	1751
OY	1739	GCATGTGTGCAAGATCTTAAGAAACGCTTCTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1798
Db	1752	GAATGTGTGCAAACTGTCAAGAAACGCTTCTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1811
OY	1799	GGTACACAGTCTTATTCACATCTGT	1858
Db	1812	GCTACACAGTCTTATTCACATCTGT	1871
OY	1859	ACAACGTGTGAGGAGCTTTTGT	1918
Db	1872	ACAACGTGTGAGGAGCTTTTGT	1931
OY	1919	CTCAGGCAAGCATTAAGGAAGACCCCTGTGAACCTCTACATGTGTGTGTGTGTGTGTGTGTGT	1978
Db	1932	CCGAGGCAAGCATTAAGGAAGACCCCTGTGAACCTCTACATGTGTGTGTGTGTGTGTGTGTGT	1991

Oy	1979	ATGGGCTCTCTCCCAACAGCGGGAAGACTGCGCTTCTGTGACTCCAGATGTTCTTTGCCAATA	2038
Dp	1992	ACGGGCTCTCTCCCGCGGCGAGAGGACTGCGCTTCCCGGCTCAGATGTTCTTGCGCTAATA	2051
Oy	2039	ACCATGACAGGAATTTGACCCGCCCAAGGTTTACCACCTGTGTGCCAGCTGAGAAAGAGA	2098
Dp	2052	ACCAAGACAGGAATTTTGACCCCTCCAAAGGTTTACCACCTGTGTGCCAGCTGAGAAAGAGA	2111
Oy	2099	ACCCCATCCCGTGTCTCTCTTTTGATGAGTGTCTACAGGGCTCTGTGCTGAAG	2158
Dp	2112	ACCCCATCCCGGTTGTCTCTCTTTGATGGAATGTCTACAGGGCTCTGTGCTGAAG	2171
Oy	2159	ACCTGGGATCTCAAGTGGACCGCTACATTGTCCGTCCAGAGTGTGTAGAGACTCCATCACGG	2218
Dp	2172	ACCTGGGATCTCAAGTGGAGCCGCTACATTGTCCGTCCAGAGTGTGTAGAGACTCCATCACGG	2231
Oy	2219	TGGGATGTGTGCGGCGCACAGGAAAAGATTCATGTACGTCCGGAGCGTCCGAGGCTCACAC	2278
Dp	2232	TGGGATGTGTGCGGCGCACAGGGAAGATTCATGTACGTCCGGAGCGTCCGAGGCTCACAC	2291
Oy	2279	ACGACATATTCAGAGAGTGGGGCCCATTCGACCTGTGTGATTGGAGGCAATCCCTCCGAATG	2338
Dp	2292	ACGACATATTCAGAGAGTGGGGCCCATTCGACCTGTGTGATTGGAGGCAATCCCTCCGAATG	2351
Oy	2339	ACCTCTCCATTTGTCAACCCCTGGCCCGCAAGGAGACTTTTATAGAGTACAGCGCGCTCTTCT	2398
Dp	2352	ACCTCTCCATTTGTCAACCCCTGGCGCAAGGAGCCCTTCACAGAGGCACTAGCGCGCTCTTCT	2411
Oy	2399	TTGAGTTCTTACCCCTCTCGCATGATGTCGGGCGCCCAAGAGGAGATGATCGGCCCTTCT	2458
Dp	2412	TTGAGTTCTTACCCCTCTCGCATGATGTCGGGCGCCCAAGAGGAGATGATCGGCCCTTCT	2471
Oy	2459	TCTGGCTCTTTGAGAATGTGTGGCGCATGTGGCGCTTGTGTACAGAGGAGCATGTCCGAT	2518
Dp	2472	TCTGGCTCTTTGAGAATGTGTGGCGCATGTGGCGCTTGTGTACAGAGGAGCATGTCCGAT	2531
Oy	2519	TTCCTGAGTGTAAACCCCGATGATGTGACCCCAAAATAATGTCTGTGTGACACAGGCGCC	2578
Dp	2532	TTCCTGAGTGTAAACCCCTGTGATGATGTATCCAAATAATGTCTGTGTGACACAGGCGCC	2591
Oy	2579	GTATCTTCTGGGTAACTCTCTGGCATGTGAAGAGGCGCTTGGCATCCACTGTAAATGATA	2638
Dp	2592	GCTACTTCTGGGTAACTCTCCGGTATGAAGAGGCGCGTGGCATCCACTGTAAATGATA	2651
Oy	2639	AGCTGGAGCTGCAGAGGTGTGTGGAGCAGCGGCAAGTATACCAGATGTGAGAAAGTAGGA	2698
Dp	2652	AGCTGGAGCTGCAGAGGTGTGTGGAGCAGCGGCAAGTATACCAGATGTGAGAAAGTAGGA	2711
Oy	2699	GCATTACCACAGGTCAAACTGTATAAGAGAGGCGAAAGACAGCATTTCCCGCTTCA	2758
Dp	2712	GCATTACCACAGGTCAAACTGCATAAAGAGGCGAAAGACAGCATTTCCCGCTTCA	2771
Oy	2759	TGSAAGAGAAAGAGACATCCTGTGTGTGTGCACATAATGSAAGAGGTGTGTGTGCTCCCG	2818
Dp	2772	TGSAATGAGAAAGAGACATCTTATGTGTGTGCACATAATGSAAGAGGTGTGTGTGCTCCAG	2831
Oy	2819	TTCGACTACAGAGCTGTCCAAATGAGCGGCTTGGCGAGGACAGACTGTGTGGCCGAT	2878
Dp	2832	TTCGACTACTGTAGCTGTCCAAATGAGCGGCTTGGCGAGGAGAGACTGTGTGGCCGCT	2891
Oy	2879	CGTGGAGCTGTCCGCTCATCCGCCACTTGTGGCTCCGCTGAGAGAAATATTTTGTCTGTG	2938
Dp	2892	CATGGACGTGTCCAGTCACTCCGCAACCTGTGTGCTCCGCTGAGAGAAATATTTTGTCTGTG	2951
Oy	2939	TGTAAAGACATGGGGCAAACTGAAAGTAGTGTATATAAAAAGTTTAAACAAACAAACAA	2998
Dp	2952	TGTAAAGACATGGGGCAAACTAGGTAGCGA --- CACAAAGTTTAAACAAACAAAC ---	3005
Oy	2999	ACAAAAACAAACAAACAAATATAAACCAACAAACGAGAGGAGCGAGAAAGT -TCAGC	3057
Dp	3006	--- AAAAAACAAACAAACAAATATAAACCAACAAACATAGAGATGTGAGAGAAAGTATCAGC	3062

OY	305E	ACCCAGAAAGAAAAAGAAATTTAAAG--AAACGACAGAGAGAGAAACGCCGAGGGC	311.5
Db	306J	ACCCAGAAAGAAAAAGAAATTTAAACAAAACACAGAGCGGAAATACCGAGAGCT	312.2
OY	311E	TTGGCGTTGGAAAAGCTTGACATCATATCTCGAAGTTTCAATGTTAACTCTAGCTCT	317.5
Db	312J	TTGCGTTCCAAAAGGTTTGACATCATCTCGAATTTTTCAAATGTTATCTTCACCTCT	318.2
OY	317E	ATCTAAAGAACAAATAGCG-CCCTCCCTCTCTCC--TCGGTCTCAGAGAGCGAA	323.1
Db	318J	ATTTTAAACAAACCAACCTCCCTCCCTCCCTCCCTCTCTCTCTCTCTCTCTCTCTCT	324.2
OY	323J	CTTTTGTCTTCTACT	329.1
Db	324J	CCCTTTATTTTCTACMCT-TTTTAGAGGGGTTTTCTGTTTTGTTGGG-TTTTGTCTCTG	330.0
OY	329J	CTGTGACTGAACACAGAGATTATTGGAGC-AAATCAGTAAACAACAAAAGTAGAAATG	335.0
Db	330J	CTGTGACTGAACACAGAGATTATTGGAGCAAAAATCAGTAAACAAAAGTAGAACATA	336.0
OY	335J	CCTTGGAGAGAAAG-----GAGAGAGCAAAATCTATATAAACTTAAATATT	340.1
Db	336J	CCTTGGAGAGAAAGTTGGAGAGAGAAAAAGAAATCTATAGAAACATATATTATG	342.0
OY	340J	GGTTTTTTTTTTTTTCCCTT-----	343.0
Db	342J	GGTTTTTTTTTTTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCTATATATTC	348.0
OY	343J	TTCTTTGGTGTCTCTTACCCTGATCAGTAGAGACACAAACA--GGAAAGATATAGGA	348.8
Db	348J	TTTTTTTTTTGTCTCTTACCCTGATCAGTAGAGACACACAGAGAGAGGAAAGAGGA	354.0
OY	348J	CCCTCGAGAGGCA-ATCTCCCTCTCCACCCCGACAGCTCAAGAGACATTTCTG	354.7
Db	354J	GACACTGAGCGGAGACATTCCTCTCCAGCACTGATGCTGTGCGACAGACATTTCTG	360.0
OY	354E	GTCTATGAAAACGAAACCCAACTACGACGAGGCGCTGAGAAACACACACACACACT	360.7
Db	360J	GTCTATGAAAACGAAACCCAACTATGACGACGAGGAGACGAGAACACACACACACACT	365.9
OY	360J	TTCTACGATTTTCAGTGCCTTACCACACAGAAACCTTGAAGAAACAGCTTTCTAGAA	366.7
Db	366J	TTCTACGATTTTCAGTGCCTTACCACACAGAAACCTTGAAGAAATCTGTTTCTAGAA	371.9
OY	366E	GGCGGTATACCTCTTGTATTACGTTTATATATATATATATATATATATATATATAT	372.7
Db	372J	GGCGGTATACCTCTTGTATTACGTTTATATATATATATATATATATATATATATAT	377.7
OY	372E	TAAAGGACTGTTAACTACTGTACATCCGACCTTCATATATGTTGCTTTCAAAACAGGA	378.7
Db	377E	TAAAGGACTGTTAACTACTGTACAAACCGACCTTCATATATGTTGCTTTCAAAACAGGA	383.6
OY	378E	GATGAGAAAGACATAGTTCCGCGCGCTCTGTGCAAAAGGTTTCAGCCAGAGATG	384.7
Db	383J	GATGAGTAAACAAATCAGCTTCCAGCTTGCCTCTGCGAAAGGTTTCACAAAGATGG	389.6
OY	384E	GGGAGAGAGGAGAGCTGGAGGGGTTTTTAAACAACTGAAGATGACCCATATACACCCC	390.7
Db	389J	AGAAAGGAGACAGCTTGAGTGGCGCGTTTCACGGGGGCTCTTCCCTTGGTTGT	395.6
OY	390E	CAC-----CCGTGCCCCATCGCTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	395.4
Db	395J	AACGAACTGAAGAGAGAGAACTTGGAGAGCAGGTTCTCTCTCTCTCTCTCTCTCTCTCT	401.6
OY	395E	GAGTGGTCCGACCTCGGGAACTGAGTGTGGATTTATCTCAACACTCGCGTGCATTAAC	401.4
Db	401J	GAGTGGTCCGAGCGCTGGACAGCTGAGATGGGATCATCTCCAGCTCATGCAATTAACC	407.6
OY	401E	CTTGAATATGAATCTAAATGACTGCTCAGAAAAATGGC-----TTGAGAAAAATT	406.8
Db	407J	TTTGAATTTTCTTAAAGAGACTCTCCCTCGAGATGAGAGAGGCTACCGAGCTTTCA	413.6
OY	406E	GTCCCTATTTTCAATGTGTCAGCCAGGTTGAAGGCGCCTTGTGGATCGAAATATTCC	412.8

Db 4137 GGCACGCTTCTCTACTTAGCAATTCGAGGGCTCTTGCTGGTGGATVACAGAACTATATCC 4196
QY 4129 AGAGTGGGGAAGTGTACCCGCCCTTAACCCNCTGAGCAATATAAAACATACAAA 4188
Db 4197 AGAGTGGGGAAGTGTACCAAA- AACCCGCTGAGCAATATAAAACATACAAA 4254
QY 4189 ATGT 4192
Db 4255 ACGT 4258
RESULT 4
AF067972
LOCUS 3005 bp mRNA linear PRI 12-FEB-2001
DEFINITION Homo sapiens DNA cytosine methyltransferase 3 alpha (DNMT3A) mRNA,
complete cds.
AF067972
AF067972.2 GI:12746531
ORIGIN
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3005)
Xie,S., Wang,Z., Okano,M., Nogami,M., Li,Y., He,W.W., Okumura,K.
and Li,E.
Cloning, expression and chromosome locations of the human DNMT3
gene family
Gene 236 (1), 87-95 (1999)
JOURNAL MEDLINE
99365304
PUBMED 10433969
REFERENCE
2 (bases 1 to 3005)
Xie,S., Okano,M. and Li,E.
Direct Submission
Submitted (25-MAY-1998) CVRC, Mass. General Hospital, 149 13th
Street, Charlestown, MA 02129, USA
3 (bases 1 to 3005)
Hata,K., Shirohzu,H., Sasaki,H. and En,L.
Direct Submission
Submitted (12-FEB-2001) CVRC, Mass. General Hospital, 149 13th
Street, Charlestown, MA 02129, USA
REMARK
COMMENT
FEATURES
On Feb 12, 2001 this sequence version replaced gi:4927369.
location/qualifiers
1..3005
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/map="p23; between D2S171 and D2S174; 48-50 cM"
1..3005
/gene="DNMT3A"
230..2968
/gene="DNMT3A"
/function="de novo DNA cytosine methyltransferase"
/codon_start=1
/product="DNA cytosine methyltransferase 3 alpha"
/protein_id="AAD33084.2"
/db_xref="GI:12746532"
/translation="MPAMPSSGPDVTSSSAEREDRRKDEOEPRKREORPEST
ARKVGRPKRHPPEVSGDTPKDPVAVSKSPMAODGASGLPGLNLEKRSQPE
EGSPAGGKGAGPAEGEAGETLPASAVNGCCTPEGKRADEAGKEKRNIES
MKMGSRGLRGLGWSLSLRORMPRTTPAGDPYVSKRDEMLAKMRKEKKA
KVLGNMAVENOGPESQKYEASPRAYOPTDPASTVATTFEPVSDGDKNATK
AGDEPEDEDSRFGIGELVWGLNGFSWPGRTVSWMTGRSRAEGTRVMMGDS
KFSVYCEKLMPLLSFCSAFQATYNNKQPMRKAIYEVLOVASSAGCLFVCHDSDE
SDTAAVEONKPMIEMALGFOPSGPGLRPEBEKNPYEVYTDMMVPEVAAADP
PPPAKPKRSTAEKPKVEIKIDERTRELVEVROKCNIDICDSCSLVNTLEHPL
FVGMCONCKKFLCAAYOYDDGVOYSCYICCGRGLYLMGNNCCGCFVECDIL
VGRGAADAIKEDPNCYMGCHKGTGTLRRRDEPMSLQMPRANHNDQEPDPKRYV
FVPAKRRKPIRVLSLDIATGLVLKLGSLQVVDYIYLSVECDSTIVGMVHQKIM
YVGDVRSYVORHIOEMGPFGLVIGSGPNDISIVNPAKRGYEGTGRGLFEPFYRLHD
ARPEGDGDPFFFLFENVVAMGVSDKRDISTRFLESNPVMIDAKVEASAHARAYFGNL

PGMRPLASTVNDKLELOECLERHRIAKFSVRTITTRNSIKOGDOHPEVENEKE
DILWCTMERVFPGPVHYTDVSNRSRLAROLLGRSVSPVIRHLFAPLKEFYACV"
BASE COUNT 674 a 846 c 978 g 507 t
ORIGIN
Query Match 57.1%; Score 2394.6; DB 9; Length 3005;
Best Local Similarity 89.7%; Pred. No. 0;
Matches 2614; Conservative 0; Mismatches 287; Indels 12; Gaps 4;
QY 69 CTGAGGCTGCACCTTTTCCGAGGCTTGACATCAGGCTGTATGTTTAAGTCTTACTCTGTG 128
Db 86 CAGAGGACCCCTTCACACAGAGGCTCAACACCGGATCTATGTTTAAGTCTTACTCTGTG 145
QY 129 CTTAAGAAAGACACGAGGACCTTCCTGTGAAGCCCTGAGCCGACGAGCCGCTGCA 188
Db 146 CTTCAAGAAAGACACGAGATTAATCTTCCCAAGGCC-AGCAGCCCCCAGC-CCGCGCA 203
QY 189 GCCCAGGCTCCGCTACTGCCA-----GCAATGCCCTGACAGCGGCCCGGGACA 241
Db 204 GCCCAGGCTCCGCTCCGCGCGCGAGATGCCCGCATGCCCTGACGCGGCCCGGGACA 263
QY 242 CCAGCAGCTCCTCTCTGAGCGGAGATGATCGAAAGAGAGAGAGAGAGAGAGAGAGAG 301
Db 264 CCAGCAGCTCTGCTGCGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 323
QY 302 ACCGAGGAG 361
Db 324 CGCTGAG 383
QY 362 GCCGAAAG 421
Db 384 GGAG 443
QY 422 CCAGCAGCTCTGAG 481
Db 444 TCTCCAACTCCCATCTCATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 503
QY 482 ACTTGAG 541
Db 504 ACTTGAG 563
QY 542 GTGGGGCCCGAG 598
Db 564 GCGGGGCCCGAG 623
QY 599 TGGAGATGCTGCTGTGTGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 658
Db 624 TGGAAATGCTGCTGACACCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 683
QY 659 AACAGACAGACCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 718
Db 684 AACAG 743
QY 719 GTGGCTGGGCTGGAGAGTCAAGCCTCCCTGACGAGCCCATGCAAGACTCACTTCACAG 778
Db 744 GTGGCTGGGCTGGAGAGTCAAGCCTCCCTGACGAGCCCATGCGGAGGCTCACTTCACAG 803
QY 779 CAGGGGACCCCTACTACTACATCAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 838
Db 804 CCGGGGACCCCTACTACTACATCAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 863
QY 839 GGGAGCTGAG 898
Db 864 GGGAGCTGAG 923
QY 899 CCTCTGAGAGTCTCAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 958
Db 924 GGGCGGGAGAGTCTCAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 983
QY 959 ACCCTGCTTCTCCAGCTGTGGCAGCAGCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1018
Db 984 ACCCGGATCCCGCCAGCTGTGCTACAGCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1043

1019 AGAATGCTACCAAGACCCAGCATGAGCTGATGAGATGCGGGGCTTTGGCA 1078
1044 AGAATGCCACCAAGACCGATGAGCGACAGTACAGAGAGAGCGGGGCTTTGGCA 1103
1079 TTGGAGAGCTGGTGGGGGAACTTGGGGCTTCCCTGGTGGCGAGCGCAATTTGT 1138
1104 TTGGAGAGCTGGTGGGGGAACTGCGGGCTTCCCTGGTGGCGAGCGCGCTTTGT 1163
1139 CTGGTGGATGACAGCGCGAGCGAGCTGAAAGGACCTGCGGGCTGATGGTGG 1198
1164 CTGGTGGATGACAGCGCGAGCGAGCTGAAAGGACCTGCGGGCTGATGGTGG 1223
1199 GAGATGGCACTTCTCAATGGTGTGTGTGAGAAAGCTATGCGCGCTGAGCTCTTGC 1258
1224 GAGATGGCACTTCTCAATGGTGTGTGTGAGAAAGCTATGCGCGCTGAGCTCTTGC 1283
1259 GTGCAATTCACAGCGCCACTACAAAGAGCCCATGTACCGCAAGCCATCTACGA 1318
1284 GTGCAATTCACAGCGCCACTACAAAGAGCCCATGTACCGCAAGCCATCTACGA 1343
1319 TCCTCCAGTGGCCAGAGCCGCTGCGGGAAGCTTTTCCAGCTTGCATGAGTGTATG 1378
1344 TCCTCCAGTGGCCAGAGCCGCTGCGGGAAGCTTTTCCAGCTTGCATGAGTGTATG 1403
1379 AAGATGACAGTGGCAAGCTGTGTGAGTGCAGAAAGAGATGATGATGGGCGCTG 1438
1404 AAGATGACAGTGGCAAGCTGTGTGAGTGCAGAAAGAGATGATGATGGGCGCTG 1463
1439 GTGCAATTCACAGCGCCACTACAAAGAGCCCATGTACCGCAAGCCATCTACGA 1498
1464 GTGCAATTCACAGCGCCACTACAAAGAGCCCATGTACCGCAAGCCATCTACGA 1523
1499 ACAAGAGATTTACACGAGATGTGGTGGAGCTGTAAGAGAGCTGTACGCGCCAGCC 1558
1524 ACAAGAGATTTACACGAGATGTGGTGGAGCTGTAAGAGAGCTGTACGCGCCAGCC 1583
1559 CACGAGCCCAAGAAAGCCCAAGAGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1618
1584 CACGAGCCCAAGAAAGCCCAAGAGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1643
1619 ATGAGCGCAAG 1678
1644 ATGAGCGCAAG 1703
1679 ACATTTGATCTCATGTGGAGAGCTCATGTGACCCCTGAGAGCCCACTTTCATTTGAG 1738
1704 ACATTTGATCTCATGTGGAGAGCTCATGTGACCCCTGAGAGCCCACTTTCATTTGAG 1763
1739 GCATGTGGCAG 1798
1764 GCATGTGGCAG 1823
1799 GGTACCACTGCTATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1858
1824 GGTACCACTGCTATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1883
1859 ACAATGCTGAG 1918
1884 ACAATGCTGAG 1943
1919 CTCAGGCAAGCATTTAAG 1978
1944 CCCAGGCAAGCATTTAAG 2003
1979 ATGGGCTGCTGCGAG 2038
2004 ATGGGCTGCTGCGAG 2063
2039 ACCATGACAGAGATTTGAG 2098
2064 ACCATGACAGAGATTTGAG 2123
2099 ACCCATGACAGAGATTTGAG 2158

2124 AGCCCATCCGGGCTGTGTCTCTTTGATGAAATGCGTACAGGGCTCTGTGTCTGAAG 2183
2159 ACCTGGGATCCAGATGAG 2218
2184 ACTTGGGATCCAGATGAG 2243
2219 TGGGATGCTGCGGAG 2278
2244 TGGGATGCTGCGGAG 2303
2279 AGAAGCATATCCAG 2338
2304 AGAAGCATATCCAG 2363
2339 ACCTTGCATTTGCAACCCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2398
2364 ACCTTGCATTTGCAACCCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2423
2399 TTGAGTTCTACCGGCTCTGATGATGCGGCGCCCAAGAGAGAGAGATGATGCGCCCTTCT 2458
2424 TTGAGTTCTACCGGCTCTGATGATGCGGCGCCCAAGAGAGAGAGATGATGCGCCCTTCT 2483
2459 TCTGGCTTTTGAAGATGTGGTGGCCATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2518
2484 TCTGGCTTTTGAAGATGTGGTGGCCATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2543
2519 TCTGGATGATTAACCGGCTGATGATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2578
2544 TCTGGATGATTAACCGGCTGATGATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2603
2579 GTTACTTGTGGGATTAACCGGCTGATGATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 2638
2604 GTTACTTGTGGGATTAACCGGCTGATGATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 2663
2639 ACCTGAGAGCTGCAAG 2698
2664 ACCTGAGAGCTGCAAG 2723
2699 CCATTACCAAG 2758
2724 CCATTACCAAG 2783
2759 TGAAG 2818
2784 TGAAG 2843
2819 TCCACTACACAG 2878
2844 TCCACTACACAG 2903
2879 CGTGGAG 2938
2904 CATGAG 2963
2939 TGAAG 2971
2964 TGAAG 2996

RESULT 5
AK025230
LOCUS
DEFINITION
AK025230 2191 bp mRNA linear PRI 29-SEP-2000
Homo sapiens cDNA: FLJ21577 fls, clone COL06724, highly similar to
AF067972 Homo sapiens DNA cytosine methyltransferase 3 alpha
(DNMT3A) mRNA.
ACCESSION
AK025230.1 GI:10437699
VERSION
oligo cloning: fls (full insert sequence).
KEYWORDS
Homo sapiens colon cDNA to mRNA, clone_11b:COL clone:COL06724.
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (sites)
AUTHORS Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
TITLE NEO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2191)
AUTHORS Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
TITLE Direct Submission
JOURNAL Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure Analysis, Human
Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
Japan (E-mail:cdna@leims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT NEO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan: cDNA full insert
sequencing: Research Association for Biotechnology: cDNA Library
construction, 5'- & 3'-end one pass sequencing: Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
FEATURES
Source Location/Qualifiers
1. 2191
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="COL06724"
/tissue_type="colon"
/clone_lib="COL"
/note="cloning vector [pME18SFL3]"
misc_feature
1. 2191
/note="highly similar to AF067972 Homo sapiens DNA
cytosine methyltransferase 3 alpha (DNMT3A) mRNA"
BASE COUNT 571 a 505 c 575 g 540 t
ORIGIN
Query Match 37.0%; Score 1551.6; DB 9; Length 2191;
Best Local Similarity 86.8%; Pred. No. 0;
Matches 1901; Conservative 0; Mismatches 224; Indels 65; Gaps 15;
OY 1615 ATTATGAGCGCCAGAGGAGCGGCTGTATGAGTGGCCGAGAGTGCAGAAACATC 1674
DB 2 ATTATGAGCGCCAGAGGAGCGGCTGTATGAGTGGCCGAGAGTGCAGAAACATC 61
OY 1675 GAGGACATTGTATCTCATGTGGGAGCGCTCAATGCAACCGAGACACCCCTCTCTT 1734
DB 62 GAGGACATTGTATCTCTGTGGGAGCGCTCAATGCAATGCAATGCAATGCAATGCAAT 121
OY 1735 GGAGGACATGTCAGAACTGTAAAGACTCTTCTTGAGTGTGCTTACAGTATGAGCAG 1794
DB 122 GGAGGACATGTCAGAACTGTAAAGACTCTTCTTGAGTGTGCTTACAGTATGAGCAG 181
OY 1795 GATGGATGACAGTCTTATGACATCTCTGTGGGGGCGGTGAAGTCTCATGTGTGG 1854
DB 182 GAGGACATGACAGTCTTATGACATCTCTGTGGGGGCGGTGAAGTCTCATGTGTGG 241
OY 1855 AACAAACAACCTGAGGTGCTTTGTTGAGTGTGAGTGTGATCTGTTGGGGGCGAGGA 1914
DB 242 AACAAACAACCTGAGGTGCTTTGTTGAGTGTGAGTGTGATCTGTTGGGGGCGAGGA 301
OY 1915 GCTGCTCAGGACGACATTAAGAGAGACCCCTGGAAGTCTTACATGTGGGGGCGAGG 1974
DB 302 GCTGCTCAGGACGACATTAAGAGAGACCCCTGGAAGTCTTACATGTGGGGGCGAGG 361
OY 1975 ACCTATGGGCTGTGCGAAGAGAGAGTGGCTTGTGACATCTCAGATGTTCTTGGC 2034
DB 362 ACCTATGGGCTGTGCGAAGAGAGAGTGGCTTGTGACATCTCAGATGTTCTTGGC 421
OY 2035 AATAACCACTGACAGGAATTTGACCCCCCAAGGTTTACCAACCTGCGGACGAGAG 2094
DB 422 AATAACCACTGACAGGAATTTGACCCCCCAAGGTTTACCAACCTGCGGACGAGAG 481
OY 2095 AGGAAGCCATCCGCTGCTCTCTTTGATGGATGCTACAGGGCTCTGCTGCTG 2154

DB 482 AGGAAGCCATCCGCTGCTCTCTTTGATGGATGCTACAGGGCTCTGCTGCTG 541
OY 2155 AAGACCTGGGACATCCAGAGTGGAGCCGTACTTGTGCTCCGAGGTGTGAGACTCCATC 2214
DB 542 AAGACCTGGGACATCCAGAGTGGAGCCGTACTTGTGCTCCGAGGTGTGAGACTCCATC 601
OY 2215 ACGGTGGGACATGGGCGGACACAGGAAAGTATGATACGTGGGGAGCGTCCGACGCTC 2274
DB 602 ACGGTGGGACATGGGCGGACACAGGAAAGTATGATACGTGGGGAGCGTCCGACGCTC 661
OY 2275 ACACAGAAAGCATATCCAGAGTGGGCGGACATTCAGTGTGATTTGAGAGCACTCCCTGC 2334
DB 662 ACACAGAAAGCATATCCAGAGTGGGCGGACATTCAGTGTGATTTGAGAGCACTCCCTGC 721
OY 2335 AATGACCTCTCCATGTCACCCCTGCGGACAGGAGACTTTATGAGGTACTGCGCGCTC 2394
DB 722 AATGACCTCTCCATGTCACCCCTGCGGACAGGAGACTTTATGAGGTACTGCGCGCTC 781
OY 2395 TTCTTTGAGTCTACCGGCTCTGATGATGATGCGGCGGACAGGAGGATGATGCGGCC 2454
DB 782 TTCTTTGAGTCTACCGGCTCTGATGATGATGCGGCGGACAGGAGGATGATGCGGCC 841
OY 2455 TTCTTTGAGTCTTTGAGAAATGTGTGGCCATGGGCGTTACTGACAAGAGGACATCTCG 2514
DB 842 TTCTTTGAGTCTTTGAGAAATGTGTGGCCATGGGCGTTACTGACAAGAGGACATCTCG 901
OY 2515 CGATTTCTTGAGTCTAACCCCGTATGATTAAGCCCAAGAAAGTGTCTGTCACACAGG 2574
DB 902 CGATTTCTTGAGTCTAACCCCGTATGATTAAGCCCAAGAAAGTGTCTGTCACACAGG 961
OY 2575 GCCCGTACTCTGCGGGTAACTTCCGCTGATGATGATGATGATGATGATGATGATGAT 2634
DB 962 GCCCGTACTCTGCGGGTAACTTCCGCTGATGATGATGATGATGATGATGATGATGAT 1021
OY 2635 GATTAAGTGTGAGTGTGAGAGTGTGTGAGAGCAGGAGCAATAGCAATGTCAGCAAGTG 2694
DB 1022 GATTAAGTGTGAGTGTGAGAGTGTGTGAGAGCAGGAGCAATAGCAATGTCAGCAAGTG 1081
OY 2695 AGGACATTTACACAGGTCATCTTAAAGCAGGAGCAAGACATTTTCCCGCTC 2754
DB 1082 AGGACATTTACACAGGTCATCTTAAAGCAGGAGCAAGACATTTTCCCGCTC 1141
OY 2755 TTCAATGAGAGAGAGAGAGATCTGTGAGCACTGAAATGAGAAAGGCTTTTGGCTTC 2814
DB 1142 TTCAATGAGAGAGAGAGAGATCTGTGAGCACTGAAATGAGAAAGGCTTTTGGCTTC 1201
OY 2815 CCCGTCACATACAGAGAGTCTCCAAACATGAGCCGCTTGGGAGGACAGACATCTGGGC 2874
DB 1202 CCCGTCACATACAGAGAGTCTCCAAACATGAGCCGCTTGGGAGGACAGACATCTGGGC 1261
OY 2875 CGATGCTGAGAGCGCGGCTATCCGACACTCTTCCGCTGAGAGAAATATTGCT 2934
DB 1262 CGATGCTGAGAGCGCGGCTATCCGACACTCTTCCGCTGAGAGAAATATTGCT 1321
OY 2935 TGTGTGTAAGGAGATGAGGAGCAATCAATGATGATGATGATGATGATGATGATGAT 2994
DB 1322 TGTGTGTAAGGAGATGAGGAGCAATCAATGATGATGATGATGATGATGATGATGAT 1377
OY 2995 ACAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAA 3053
DB 1378 ACAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCA 1432
OY 3054 CAGACCCAG 3111
DB 1433 CAGACCCAG 1492
OY 3112 GGGCTTGGCTTGGAG 3171
DB 1493 GGGCTTGGCTTGGAG 1552
OY 3172 TCGTATCTTAAAGCAAAATAGG--CCCTCCCTTCTTCCCTTCTTCCCTTCTTCCCT 3227

* 9636 11346: contig of 1711 bp in length
* 11347 11446: gap of 100 bp
* 11447 13984: contig of 2538 bp in length
* 13985 14084: gap of 100 bp
* 14085 17045: contig of 2961 bp in length
* 17046 17145: gap of 100 bp
* 17146 34039: contig of 16894 bp in length
* 34040 34139: gap of 100 bp
* 34140 49222: contig of 15083 bp in length
* 49223 49322: gap of 100 bp
* 49323 67306: contig of 17984 bp in length
* 67307 67406: gap of 100 bp
* 67407 110848: contig of 43442 bp in length
* 110849 110948: gap of 100 bp
* 110949 147393: contig of 36445 bp in length
* 147394 147493: gap of 100 bp
* 147494 180251: contig of 32758 bp in length
* 180252 180351: gap of 100 bp
* 180352 204534: contig of 24183 bp in length.
Location/Qualifiers
1. 204534
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP23-9D1"
/clone_id="RP23-9D1"
1. 9535
/note="assembly_fragment"
clone_end:SP6
vector_side:left"
9636. 11346
/note="assembly_fragment"
11447. 13984
/note="assembly_fragment"
14085. 17045
/note="assembly_fragment"
17146. 34039
/note="assembly_fragment"
34140. 49222
/note="assembly_fragment"
49323. 67306
/note="assembly_fragment"
67407. 110848
/note="assembly_fragment"
110949. 147393
/note="assembly_fragment"
147494. 180251
/note="assembly_fragment"
180352. 204534
/note="assembly_fragment"
clone_end:T7
vector_side:right"

FEATURES

source

misc_feature
1. 9535
/note="assembly_fragment"
clone_end:SP6
vector_side:left"
9636. 11346
/note="assembly_fragment"
11447. 13984
/note="assembly_fragment"
14085. 17045
/note="assembly_fragment"
17146. 34039
/note="assembly_fragment"
34140. 49222
/note="assembly_fragment"
49323. 67306
/note="assembly_fragment"
67407. 110848
/note="assembly_fragment"
110949. 147393
/note="assembly_fragment"
147494. 180251
/note="assembly_fragment"
180352. 204534
/note="assembly_fragment"
clone_end:T7
vector_side:right"

BASE COUNT 52148 a 49228 c 50043 g 52097 t 1018 others
ORIGIN

Query Match 32.7%; Score 1369.4; DB 2: Length 204534;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1392: Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 2798 AAGGGTGTGGCTTCCCGTCACCTACACAGACGCTCCACATGAGCCGCTGGCGA 2857
Db 162698 ACAGGCTGTGGCTTCCCGTCACCTACACAGACGCTCCACATGAGCCGCTGGCGA 162757
QY 2858 GGCAGACGCTGGCGGAGTCGAGAGCGTGGGTCATCGGCACCTCTTCGCTCCG 2917
Db 162758 GGCAGACGCTGGCGGAGTCGAGAGCGTGGGTCATCGGCACCTCTTCGCTCCG 162817
QY 2918 TGAAGAAATATTTGCTGTCTGTAAAGGACATGGGGCAAACTGAAGTAGTAGATTA 2977
Db 162818 TGAAGAAATATTTGCTGTCTGTAAAGGACATGGGGCAAACTGAAGTAGTAGATTA 162877
QY 2978 AAAAGTTAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 3037
Db 162878 AAAAGTTAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 162937

QY 3038 AGGACGAGAAAAAGTTGACGACCCGAGAGAAAAAGGAATTTAAGCAACACAGAGG 3097
Db 162938 AGGACGAGAAAAAGTTGACGACCCGAGAGAAAAAGGAATTTAAGCAACACAGAGG 162997
QY 3098 AGAAAAAGCGCGAGGGCTTGGCCTTCCAAAGAGGTTGAGCATCATCTCGAGTTTCA 3157
Db 162998 AGAAAAAGCGCGAGGGCTTGGCCTTCCAAAGAGGTTGAGCATCATCTCGAGTTTCA 163057
QY 3158 ATGTTAACTTCAGTCTATCTAATAAAGCAAAATAGCCCCCTCCCTTCTCCCTCCG 3217
Db 163058 ATGTTAACTTCAGTCTATCTAATAAAGCAAAATAGCCCCCTCCCTTCTCCCTCCG 163117
QY 3218 TCTAGAGGGGAACTTTTGTTCACCTTTTTCAGAGGGGTTTCTGTTGTTGG 3277
Db 163118 TCTAGAGGGGAACTTTTGTTCACCTTTTTCAGAGGGGTTTCTGTTGTTGG 163177
QY 3278 GTTTTGTCTTCCTGCTGACTGAAACAGAGAGTTATTCGACAAATCAGTAACACA 3337
Db 163178 GTTTTGTCTTCCTGCTGACTGAAACAGAGAGTTATTCGACAAATCAGTAACACA 163237
QY 3338 AAAAGTGAATGCTTGGAGAGAGAAAGGAGAGAGGAAATTTCTATTAATACTTAAA 3397
Db 163238 AAAAGTGAATGCTTGGAGAGAGAAAGGAGAGAGGAAATTTCTATTAATACTTAAA 163297
QY 3398 TATTGTTTTTTTTTTTTTTTCTTCTATATATCTCTGTTGGTGTCTAGCTGATCA 3457
Db 163298 TATTGG-TTTTTTTTTTTTTTTTCTTCTATATATCTCTGTTGGTGTCTAGCTGATCA 163356
QY 3458 GATAGAGACACAAACAGAGAAATAGAGACCTCGAGAGCAGAGTCTCTCCACAC 3517
Db 163357 GATAGAGACACAAACAGAGAAATAGAGACCTCGAGAGCAGAGTCTCTCCACAC 163416
QY 3518 CCCCAGACAGTCTAAGACACCATTCCTGCTATGCAAAACAAACCAACCACTAGACCA 3577
Db 163417 CCCCAGACAGTCTAAGACACCATTCCTGCTATGCAAAACAAACCAACCACTAGACCA 163476
QY 3578 GGGCGTGAAGAACACACACACAGACAC-TTTTACAGATTTTCAGGTGCTACACAC 3636
Db 163477 GGGCGTGAAGAACACACACACAGACAC-TTTTACAGATTTTCAGGTGCTACACAC 163536
QY 3637 AGGAACCTTGAGAAAAACAGTTTCTAGAGCCGCTGTTACCTCTGTTTACAGTTTAT 3696
Db 163537 AGGAACCTTGAGAAAAACAGTTTCTAGAGCCGCTGTTACCTCTGTTTACAGTTTAT 163596
QY 3697 ATATATATGATATGATATGATATATATATATATATATATATATATATATATATAT 3756
Db 163597 ATATATATGATATGATATGATATATATATATATATATATATATATATATATATAT 163656
QY 3757 CGACTTCATATGCTGCTTCAAAAACAGCGAGTGAAGCAATCAGTTCGCGCTGG 3816
Db 163657 CGACTTCATATGCTGCTTCAAAAACAGCGAGTGAAGCAATCAGTTCGCGCTGG 163716
QY 3817 CCGTCTGTGCAAAAGGTTTACGCCOAGGATGGGAGAGAGCAGCTGAGGGGGTTT 3876
Db 163717 CCGTCTGTGCAAAAGGTTTACGCCOAGGATGGGAGAGAGCAGCTGAGGGGGTTT 163776
QY 3877 AACAACTGAAGATGACCCATATCACCCTCCCTGCGCCCATGCTTACCTCACCTG 3936
Db 163777 AACAACTGAAGATGACCCATATCACCCTCCCTGCGCCCATGCTTACCTCACCTG 163836
QY 3937 CCAAAAAGGCGCTAGCTGAGGTGGTGGACCTGGGAACTGATGTGAATTTATCC 3996
Db 163837 CCAAAAAGGCGCTAGCTGAGGTGGTGGACCTGGGAACTGATGTGAATTTATCC 163896
QY 3997 AGACTCGGTCATTAACCTTGAATATGAATCTTAATTAAGTACTGCTCAGAAAAATGCT 4056
Db 163897 AGACTCGGTCATTAACCTTGAATATGAATCTTAATTAAGTACTGCTCAGAAAAATGCT 163956
QY 4057 TGAGAAAAATTTGTCGATTTTGAATTCGTCAGCCAGTGAAGGCGCTTGGGAT 4116
Db 163957 TGAGAAAAATTTGTCGATTTTGAATTCGTCAGCCAGTGAAGGCGCTTGGGAT 164016

Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
Li, J., Li, Z., Lichtarge, O., Lien, C., Liu, C., Liu, W., Louisse, H.,
Lozano, R., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapa, P., Martin, R., Martindale, A., Martinez, E.,
Masssey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M.,
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokkenwo, S., Ogih, M., Okunnu, G.,
Oregany, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Ploken, R., Primus, E., Pu, L., Qullis, M., Ren, Y.,
Rivers, M., Rojas, A., Rotjokan, I., Rolfe, M., Ruiz, S., Savary, G.,
Scherer, S., Scott, G., Shen, H., Shooshbari, N., Sisson, I.,
Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Swalek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usmani, K., Vasquez, L., Vera, V., Villalob, D., Vinas, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleczek, R., Woodcock, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorilla, S., Nelson, D.,
Wetstock, C., and Gibbs, R.

Unpublished
2 (bases 1 to 123936)

Direct Submission
Submitted (13-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 123936)

Worley, K.C.

Direct Submission
Submitted (23-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 18, 2002 this sequence version replaced gi:20531783.

Genome Center

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
Project Information
Center project name: GWL
Center clone name: CH230-282D16
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye 1.00% of reads
Assembly program: Phrap, version 0.990329
Consensus quality: 9597 bases at least Q40
Consensus quality: 99223 bases at least Q30
Consensus quality: 101036 bases at least Q20

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
NOTE: This is a 'working draft' sequence. It currently
consists of 38 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1 1009: contig of 1009 bp in length
* 1010 1109: gap of unknown length
* 1110 2113: contig of 1004 bp in length
* 2114 3233: gap of unknown length
* 3234 3333: contig of 1020 bp in length
* 3334 4442: gap of unknown length
* 4443 4542: contig of 1109 bp in length
* 4543 6202: gap of unknown length
* 6203 6302: contig of 1660 bp in length
* 6303 7464: gap of unknown length
* 7464: contig of 1162 bp in length

7465 7564: gap of unknown length
* 7565 8772: contig of 1208 bp in length
* 8773 8873: gap of unknown length
* 8874 10371: contig of 1499 bp in length
* 10372 10472: gap of unknown length
* 10473 11714: contig of 1243 bp in length
* 11715 11815: gap of unknown length
* 11816 12924: contig of 1110 bp in length
* 12925 13025: gap of unknown length
* 13026 14347: contig of 1323 bp in length
* 14348 14448: gap of unknown length
* 14449 15522: contig of 1205 bp in length
* 15523 15752: gap of unknown length
* 15753 17796: contig of 2044 bp in length
* 17797 17966: gap of unknown length
* 17967 18937: contig of 1041 bp in length
* 18938 19037: gap of unknown length
* 19038 21543: contig of 2506 bp in length
* 21544 24540: gap of unknown length
* 24541 24640: contig of 2897 bp in length
* 24641 26112: gap of unknown length
* 26113 28112: contig of 3472 bp in length
* 28113 31101: contig of 2869 bp in length
* 31102 31202: gap of unknown length
* 31203 34115: contig of 3214 bp in length
* 34116 34515: gap of unknown length
* 34516 37353: contig of 2838 bp in length
* 37354 37454: gap of unknown length
* 37455 40938: contig of 3485 bp in length
* 40939 44038: gap of unknown length
* 44039 44331: contig of 3293 bp in length
* 44332 44431: gap of unknown length
* 44432 48035: contig of 3604 bp in length
* 48036 48135: gap of unknown length
* 48136 51181: contig of 3046 bp in length
* 51182 53561: gap of unknown length
* 53562 55461: contig of 4080 bp in length
* 55462 59680: gap of unknown length
* 59681 63525: contig of 4119 bp in length
* 63526 63625: gap of unknown length
* 63626 67434: contig of 3809 bp in length
* 67435 72194: gap of unknown length
* 72195 72294: contig of 4660 bp in length
* 72295 77778: gap of unknown length
* 77779 81349: gap of unknown length
* 81350 81449: gap of unknown length
* 81450 86396: contig of 4947 bp in length
* 86397 91224: gap of unknown length
* 91225 97769: contig of 4728 bp in length
* 97770 97869: gap of unknown length
* 97870 105670: contig of 6445 bp in length
* 105671 105770: gap of unknown length
* 105771 111368: gap of 7801 bp in length
* 111369 111468: gap of unknown length
* 111469 116680: contig of 5598 bp in length
* 116681 116780: gap of unknown length
* 116781 123936: contig of 5212 bp in length
* 123937 123938: gap of unknown length
* 123939 123940: contig of 7156 bp in length

Location/Qualifiers
1. 123936
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="CH230-282D16"

BASE COUNT 35904 a 21413 c 21468 g 36486 t 8665 others
ORIGIN
Query Match 31.8% Score 1331.8; DB 2; Length 123936;

Best Local Similarity 85.6%; Pred. No. 9,2e-307;
Matches 1485; Conservative 0; Mismatches 217; Indels 33; Gaps 5;

Qy	843	GGCTGAGAAAGAAACCAAGTAAATTCAGTAATATATGCTGTGGAGAGAACAGCGCTC	902	Db	85801	AACCGCTGACAGTCTTTTGTGTGAGTGTGTGATCTCTTGTGTGGGCCAGGGCTGCC	85860
Db	84794	GGTTGAGAAAGAAACCAAGTAAATTCAGTAATATGCTGTGGAGAGAACAGCGCTC	84853	Qy	1921	CAGCAGCATTAAGAGAAAGACCCCTGGAACCTGCTACATGTGTGGGCAATAAGGCACCTAT	1980
Qy	903	TGAGAGCTCAGAAAGGTGGAGAGAGCCCTCTGCTGTGTGAGAGCCAGGAGCC	962	Db	85861	CAACGAGCATTAAGAGAAAGACCCCTGGAACCTGCTACATGTGTGGGCAATAAGGCACCTAT	85902
Db	84854	TGGGAGTCTCAGAAAGGTGGAGAGAGCCCTCTGCTGTGTGAGAGCCAGGAGCC	84912	Qy	1981	GGGCTGCTGAGAAAGGAGAGTGGCTCTGCTGAGCTCCAGATGTTCTTGGCAATTAAC	2040
Qy	963	TGCTTCTCCAGCTGTGGCCACCCCTGAGCCAGTAAGAGAGGAGTGTGGGAGAGAA	1022	Db	85903	GAGCTGCTGAGAGAGAGAGTGGCTCTGCTGAGCTCCAGATGTTCTTGGCAATTAAC	85962
Db	84913	TGCTTCTCCAGCTGTGGCCACCCCTGAGCCAGTAAGAGAGGAGTGTGGGAGAGAA	84972	Qy	2041	CATGACAGGAATTTGACCCCCCAAGGTTTACCACCTGTGTCCAGCTGAGAAGAGAG	2100
Qy	1023	TGCTTCTCCAGCTGTGGCCACCCCTGAGCCAGTAAGAGAGGAGTGTGGGAGAGAA	1082	Db	85963	CACGACAGGAATTTGACCCCCCAAGGTTTACCACCTGTGTCCAGCTGAGAAGAGAG	86022
Db	84973	TGCTTCTCCAGCTGTGGCCACCCCTGAGCCAGTAAGAGAGGAGTGTGGGAGAGAA	85032	Qy	2101	CCCATCCGCGCTGCTCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2160
Qy	1083	AGACTGTGTGTGGGAGAACTTGGGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG	1142	Db	86023	CCCATCCGCGCTGCTCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	86082
Db	85033	AGACTGTGTGTGGGAGAACTTGGGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG	85092	Qy	2161	CTGGGATCCAAAGTGGAGACCGCTACATTCCTCCAGAGTGTGTGTGTGTGTGTGTGT	2220
Qy	1143	GTGATGACAGAGCCGAGCCAGCAGTGAAGGCACTGCTGGGTCAATGTGTGTGTGTGT	1202	Db	86083	CTGGGATCCAAAGTGGAGACCGCTACATTCCTCCAGAGTGTGTGTGTGTGTGTGTGT	86142
Db	85093	GTGATGACAGAGCCGAGCCAGCAGTGAAGGCACTGCTGGGTCAATGTGTGTGTGTGT	85152	Qy	2221	GGCATGT	2280
Qy	1203	TGCGAAGTCTCAATGT	1262	Db	86143	GGCATGT	86202
Db	85153	TGCGAAGTCTCAATGT	85208	Qy	2281	AAGCATATCCAGAGTGGGGCCCATTCACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2340
Qy	1263	ATTCACACAGGCGCCTACACAGAGAGCCATGACCGCAAGGCAATCTAGAGAGTCT	1322	Db	86203	AAGCATATCCAGAGTGGGGCCCATTCACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	86262
Db	85209	ATTCACACAGGCGCCTACACAGAGAGCCATGACCGCAAGGCAATCTAGAGAGTCT	85268	Qy	2341	CTCTCATTTGCAACCCCTGCGGCAAGGAGCTTTATGAGGCTTACCTGCGGCTCTCTCT	2400
Qy	1323	CCAGGTGGCCAGAGCCGTGCGGAGAGCTTTTCCAGCTTCCATGACAGTGTGTGTGT	1380	Db	86263	CTCTCATTTGCAACCCCTGCGGCAAGGAGCTTTATGAGGCTTACCTGCGGCTCTCTCT	86322
Db	85269	CCAGGTGGCCAGAGCCGTGCGGAGAGCTTTTCCAGCTTCCATGACAGTGTGTGTGT	85328	Qy	2401	GAGTTCTACGCGCTCTGCTCATGATGTGCGGCGCCCAAGAGAGATGATGCGCCCTCTCT	2460
Qy	1381	AGTACAGTGTGCAAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1440	Db	86323	GAGTTCTACGCGCTCTGCTCATGATGTGCGGCGCCCAAGAGAGATGATGCGCCCTCTCT	86382
Db	85329	AGTACAGTGTGCAAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	85388	Qy	2461	TGGCTCTTGTGAATGT	2520
Qy	1441	GGCTTCCAGCCCTGCGGCTCTAAGGGCTGTGAGCCACAGAAAGAGAAATGCTTAC	1500	Db	86383	TGGCTCTTGTGAATGT	86442
Db	85389	GGCTTCCAGCCCTGCGGCTCTAAGGGCTGTGAGCCACAGAAAGAGAAATGCTTAC	85448	Qy	2521	CTTGAGCTAACCCCGTGATGTGACGCCCAAGAGAGTGTGTGTGTGTGTGTGTGTGTGT	2575
Qy	1501	AAGGAAGTTTACACCGCATGTGGGTGGAGCCCTGAGAGGCTTGTACGCCCAACCCCA	1560	Db	86443	NN	86497
Db	85449	AAGGAAGTTTACACCGCATGTGGGTGGAGCCCTGAGAGGCTTGTACGCCCAACCCCA	85508	Qy	2575	CTTGAGCTAACCCCGTGATGTGACGCCCAAGAGAGTGTGTGTGTGTGTGTGTGTGTGT	2575
Qy	1561	CCAGCCAGAAACCCAG	1620	Db	2575	CTTGAGCTAACCCCGTGATGTGACGCCCAAGAGAGTGTGTGTGTGTGTGTGTGTGTGT	2575
Db	85509	CCAGCCAGAAACCCAG	85560	Qy	2575	CTTGAGCTAACCCCGTGATGTGACGCCCAAGAGAGTGTGTGTGTGTGTGTGTGTGTGT	2575
Qy	1621	GAGGGACAAAGGAGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1680	Db	2575	CTTGAGCTAACCCCGTGATGTGACGCCCAAGAGAGTGTGTGTGTGTGTGTGTGTGTGT	2575
Db	85561	GAGGGACAAAGGAGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	85620	Qy	2575	CTTGAGCTAACCCCGTGATGTGACGCCCAAGAGAGTGTGTGTGTGTGTGTGTGTGTGT	2575
Qy	1681	ATTGTATCTCATGTGGAGCCCTCAATGTCAACCCCTGAGAGCCCACTTCTTATGAGAG	1740	Db	2575	CTTGAGCTAACCCCGTGATGTGACGCCCAAGAGAGTGTGTGTGTGTGTGTGTGTGTGT	2575
Db	85621	ATTGTATCTCATGTGGAGCCCTCAATGTCAACCCCTGAGAGCCCACTTCTTATGAGAG	85680	Qy	2575	CTTGAGCTAACCCCGTGATGTGACGCCCAAGAGAGTGTGTGTGTGTGTGTGTGTGTGT	2575
Qy	1741	ATGTGCGCAAGCTGAGAACTGCTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1800	Db	2575	CTTGAGCTAACCCCGTGATGTGACGCCCAAGAGAGTGTGTGTGTGTGTGTGTGTGTGT	2575
Db	85681	ATGTGCGCAAGCTGAGAACTGCTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	85740	Qy	2575	CTTGAGCTAACCCCGTGATGTGACGCCCAAGAGAGTGTGTGTGTGTGTGTGTGTGTGT	2575
Qy	1801	TACAGTCTTCTGTACCATGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1860	Db	2575	CTTGAGCTAACCCCGTGATGTGACGCCCAAGAGAGTGTGTGTGTGTGTGTGTGTGTGT	2575
Db	85741	TACAGTCTTCTGTACCATGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	85800	Qy	2575	CTTGAGCTAACCCCGTGATGTGACGCCCAAGAGAGTGTGTGTGTGTGTGTGTGTGTGT	2575
Qy	1861	AACGTCTGACAGTCTTGT	1920	Db	2575	CTTGAGCTAACCCCGTGATGTGACGCCCAAGAGAGTGTGTGTGTGTGTGTGTGTGTGT	2575

REFERENCE

1 (bases 1 to 176697)
Murny, D.M., Adams, C., Adio-Oduola, B., Ali-Isman, F.R., Allen, C., Alsbrooks, S.L., Amaral, H.C., Are, J.R., Ayala, M., Banks, T., Bardar, J., Bencon, J., Blin, K., Blumenthal, K., Bonin, D., Bouck, J., Bowle, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P., Bunay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhury, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathore, S.R., David, R., Delaney, K.R., Delgado, O., Dem, A.L., Ding, Y., Dinn, H.H., Douthett, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Eberhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,

Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J. H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Hayak, P., Hayes, A., Hernandez, J., Hernandez, O., Hodges, A., Hogues, M., Hollaway, C., Hollins, B., Homsi, F., Howard, S., Huber, J., Hulik, S., Hume, J., Jackson, L. E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Jordan, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvan, J., Kovar, C., Kratoch, J., Kuresh, A., Landry, N., Leal, B., Lewis, L. C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lien, C., Liu, J., Liu, W., Louise, H., Lozano, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapa, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhinney, E., McLeod, M. P., Meador, M., Mei, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokwenwo, S., Ogih, M., Okunolu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Picken, R., Primus, E., Pu, L. L., Quiles, M., Ren, Y., Rivers, M., Rojas, A., Rojebokan, I., Rolfe, M., Ruiz, S., Saverly, G., Scherer, S., Scott, G., Shen, H., Shooshitari, N., Sisson, I., Sodergren, E., Sonaik, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Umanal, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, O., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G., and Gibbs, R.

Unpublished
2 (bases 1 to 176697)
Morley, K. C.
Direct Submission
Submitted (22-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 176697)
Worley, K. C.
Direct Submission
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Jul 12, 2002 this sequence version replaced g1.20303171.

Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
Project Information
Center project name: GRAT
Center clone name: CH230-28122
----- Summary Statistics
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye 100% of reads
Assembly program: Phrap version 0.990329
Consensus quality: 126371 bases at least Q40
Consensus quality: 126371 bases at least Q30
Consensus quality: 130428 bases at least Q20

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_dirft_data.html).
NOTE: This is a 'working draft' sequence. It currently
consists of 65 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1 1415: contig of 1415 bp in length
* 1416 1515: gap of unknown length
* 1516 2865: contig of 1450 bp in length
* 2866 3065: gap of unknown length

3066 4328: contig of 1163 bp in length
* 4329 4385: contig of 1057 bp in length
* 4386 5485: gap of unknown length
* 5486 6570: contig of 1185 bp in length
* 6571 7813: contig of 1043 bp in length
* 7814 7914: gap of unknown length
* 7915 9674: gap of 1760 bp in length
* 9675 9774: gap of unknown length
* 9775 11471: contig of 1697 bp in length
* 11472 11570: gap of unknown length
* 11571 13163: contig of 1592 bp in length
* 13164 13263: gap of unknown length
* 13264 14505: contig of 1242 bp in length
* 14506 14604: gap of unknown length
* 14605 15975: contig of 1370 bp in length
* 15976 16074: gap of unknown length
* 16075 17397: contig of 1322 bp in length
* 17398 17496: gap of unknown length
* 17497 19323: contig of 1827 bp in length
* 19324 19424: gap of unknown length
* 19425 21042: contig of 1618 bp in length
* 21043 21442: gap of unknown length
* 21443 22419: contig of 1278 bp in length
* 22420 22520: gap of unknown length
* 22521 23943: contig of 1423 bp in length
* 23944 24043: gap of unknown length
* 24044 25334: contig of 1291 bp in length
* 25335 25734: gap of unknown length
* 25735 26330: contig of 1057 bp in length
* 26331 26830: gap of unknown length
* 26831 28270: contig of 1640 bp in length
* 28271 29570: gap of unknown length
* 29571 30070: contig of 1600 bp in length
* 30071 32265: gap of unknown length
* 32266 32365: contig of 2195 bp in length
* 32366 34179: gap of unknown length
* 34180 34279: contig of 1814 bp in length
* 34280 36392: gap of unknown length
* 36393 36492: contig of 2113 bp in length
* 36493 38552: gap of unknown length
* 38553 38652: contig of 2060 bp in length
* 38653 40932: gap of unknown length
* 40933 41032: contig of 2280 bp in length
* 41033 43021: gap of unknown length
* 43022 43121: contig of 1989 bp in length
* 43122 44977: gap of unknown length
* 44978 45077: contig of 1856 bp in length
* 45079 47358: gap of unknown length
* 47359 47458: contig of 2281 bp in length
* 47459 49554: gap of unknown length
* 49555 49654: contig of 2096 bp in length
* 49656 51876: gap of unknown length
* 51877 51976: contig of 2222 bp in length
* 51977 53794: gap of unknown length
* 53795 53894: contig of 1818 bp in length
* 53895 55634: gap of unknown length
* 55635 55734: gap of unknown length
* 55735 57445: contig of 1711 bp in length
* 57446 57545: gap of unknown length
* 57546 60210: contig of 2665 bp in length
* 60211 60310: gap of unknown length
* 60311 62174: gap of unknown length
* 62175 62275: contig of 1864 bp in length
* 62276 64286: gap of unknown length
* 64287 64385: contig of 2011 bp in length
* 64386 65764: gap of unknown length
* 65765 66864: gap of 2379 bp in length
* 66865 69398: gap of unknown length
* 69399 71451: contig of 2534 bp in length
* 71452 71451: contig of 1953 bp in length

*	71552	gap of unknown length
*	71552	contig of 2834 bp in length
*	74406	gap of unknown length
*	74506	contig of 2081 bp in length
*	75587	75586: gap of unknown length
*	76687	80397: contig of 3711 bp in length
*	80398	gap of unknown length
*	80498	82399: contig of 1802 bp in length
*	82300	gap of unknown length
*	82400	84796: contig of 2397 bp in length
*	84797	84896: gap of unknown length
*	84897	85681: contig of 2085 bp in length
*	86982	87081: gap of unknown length
*	87082	89794: contig of 2713 bp in length
*	89795	89984: gap of unknown length
*	89895	91596: contig of 1802 bp in length
*	91697	91966: gap of unknown length
*	91797	94458: contig of 2662 bp in length
*	94459	94558: gap of unknown length
*	94559	98352: contig of 3794 bp in length
*	98353	98452: gap of unknown length
*	98453	101134: contig of 2662 bp in length
*	101135	101334: gap of unknown length
*	101235	104688: contig of 3334 bp in length
*	104589	104688: gap of unknown length
*	104689	107761: contig of 3073 bp in length
*	107762	107861: gap of unknown length
*	107862	111340: contig of 3479 bp in length
*	111341	111574: gap of unknown length
*	111441	111574: contig of 3934 bp in length

Query Match	27.98;	Score 1171.2;	DB 2;	Length 176697;
Best Local Similarity	82.7%;	Pred. No. 2.3e-268;		
Matches 1317;	Conservative	0;	Mismatches 248;	Indels 27;
				Gaps 3;

OY	124	TTGAGCTCCTCTCAGTGCATTCCACAGGCCACTCAACAAGAAGCCCATGTACCGA	1303
Db	24084	TTGTTTCTTTAGCAATPAACAATTTCTAGCTGAAGTAATTANNNNNNNNNNNNNNNN	24025
OY	1304	AAGCATTCTAGGAATCTCTCAGGTGGCCAGAGCCGTCCGGAGAGCTGTTCCAGCTT	1363
Db	24024	NN	23965
OY	1364	GCCATGACATGATGAAAGTGACAGTGGCAAGCTGTGTAAGTCAGAACAAGCAGATGA	1423
Db	23964	NN	23906
OY	1424	TTGATATGGGCTTCGGTGCTTCACGCTTCGGGTCTCTAAGGGCCCTGGAGCCACAGAAG	1483
Db	23905	TTGATGTGGTCTTTGGCGGCTTCCAGACCTCTGTGTCCAAGGCGCTCTAGCCACCTAAG	23846
OY	1484	AAGAGAAGAAVCTTTACAAGAAATTACACCGACATGTGGGTGAGGCTGAAGCAGCTG	1543
Db	23845	AGGGAACATCTTTACAAGGAAGTTTACACCGACATGTGGGTGAGGCTGAAGCAGCTG	23786
OY	1544	CTTACGGCCCAACCCCAACGAACCCGAAGAGCCACACAGAAACCTTAAG	1603
Db	23785	CTTATTTCCCGCCCAACGACGACCAAGAAACCCGAAGAGCAC-----AACCCTAAG	23734
OY	1604	TCAAGGAGATCATTTGATGAGCGCACAGAGAGCGCTGTGATGAGTGCGCCAGAAAT	1663
Db	23733	TCAAGGAGATCATTTGATGAGCACGCAAGATTAAGACTCAATGATGAGTGTGCCCAAAAGT	23674
OY	1664	GCAGAAACATCGAGAGCATTTGTATCTCATGTGGAGCTCAATGTCAACCTTGGAGCACC	1723
Db	23673	GCCAAAGCATCGAGAGCATTTGTATCTCATGTGGAGGCTCAATGTCAACCTTGGAGCACC	23614
OY	1724	CACCTTCATTTGAGAGCATGTGCCAACAACGTGAAGAACTCTTGGAGTGGCTTACC	1783
Db	23613	CACCTTCATTTGAGAGCATGTGCCAACAACGTGAAGAACTCTTGGAGTGTGCTTACC	23554
OY	1784	AGTATGACGAGATGGGTACCACTCTTAATTCACCACTGTCTGTGGGGGCGTGAAGTGC	1843

Dd	23553	AATACGATGACGATGTGTACCAGTCTCTTGTAACATCTCTGTTGGGGCCGAAGTC	23496
Oy	1844	TCATGTGTGGGAACAACAAGTGCAGAGTCTTTGTGTGCAGATGTGTGATCTCTGG	1903
Dd	23493	TCATGTGTGGGAACAACAAGCGGTGCAGAGTCTTTGTGTGCAGATGTGTGATCTCTGG	23433
Oy	1904	TGGGGCCAAGAGCTGCTCAGGCAGCCATTAAAGAAACCCCTGGAAGTGTACTATGTGC	1963
Dd	23433	TGGGGCCAAGAGGCGTGCAGCAAGCCATTAAAGAAAGCCCTTAAGTGTACTATGTG --	23376
Oy	1964	GGCATAAGGGCACCTATGTGGCTGTGTGGAAGAGGCAAGTGGCTTCTCGACTCCAGA	2023
Dd	23375	-----TGACTGTGTGGGGAAGAGAGACTGGCTTCTCGACTCCACA	23333
Oy	2024	TGTTCTTTGCCAATTAACCATGACCAGGAATTTGACCCCCCAAAGGTTTACCACCTGTGC	2083
Dd	23331	TGTTCTTTGCCAATTAACCATGACCAGGAATTTGACCCCCCAAAGGTTTACCACCTGTGC	23272
Oy	2084	CAGGTGGAAGAGAACCCATCCGCTGCTGTCTCTTTGATGGATTTGTACAGAGGC	2143
Dd	23271	CAGGTGGAAGAGGAACCCATCCAGTGTATCTTATTTGATGGATTTGTACAGAGGC	23212
Oy	2144	TCCGTGTGGAAGAGACCTGGGCAATCCAAATGGAACGCTACATTTGCTCCGAGTGTGTG	2203
Dd	23211	TCCGTGTGGAAGAGACCTGGGCAATCCAAATGGAACGCTACATTTGCTCCGAGTGTGTG	23155
Oy	2204	AGGACTCCATACCGTGTGGCATGTGTGCGCACCAAGGAAGATCATGTACGTGGGACG	2263
Dd	23151	AGGACTCCATACCGTGTGGCATGTGTGCGCACCAAGGAAGATCATGTACGTGGGACG	23092
Oy	2264	TCCCACAGCTCACACGAAGCATATCCAGAGATGGGGCCATTGCGACTGCTGATTTGGAG	2323
Dd	23091	TCTTCACAGCTCACACGAAGCATATCCAGAGATGGGGCCATTGCGACTGCTGATTTGGG	23033
Oy	2324	GCAGTCCCTGCATGACCTCTCCATTGTCAACCTGCGCGCAAGGACTTTATGAGGGTA	2383
Dd	23031	GCAGTCCCTGCATGACCTCTCCATTGTCAACCTGCGCGCAAGGACTTTATGAGGGTA	22972
Oy	2384	CTGGCCGCGCTCTTTGAGTTCTACCGCTCCGCGCATGTATGGCGGCCCAAGAGGGAG	2443
Dd	22971	CTGGCAGCGCTCTTTGAGTTCTACCACTCTCTGCATGATGCGCGGCCCAAGAGGGAG	22911
Oy	2444	ATGATCGACCCCTTCTCTGAGCTCTTTGAGAATGTGTGTGCCATGAGGCGTTAGTACAAGA	2503
Dd	22911	AGATCGACCCCTTCTCTGAGCTCTTTGAGAAGGAGTGTGTGCCATGAGGCGTTAGTACAAGA	22855
Oy	2504	GGGACATCTGGCGATTTCTTGAGTCTPACCCCGTGATGATTTGACGCCCAAGAGTGTCTG	2563
Dd	22851	GGGACATCTGGCGATTTCTTGAGTCTPACCCCGTGATGATTTGACTCCAAAAGTGTCTG	22792
Oy	2564	CTGCACACAGGGCCCGTTACTTCTGGGGTAACTTCTCTGSCATGAACAGGCGTTTGGCAT	2628
Dd	22791	CTGCACACAGGGTCCGTACTTCTGGGGTAACTTCTCTGSCATGAACAGGCGTTTGGCAT	22733
Oy	2624	CCACTGTGAATGATTAAGCTGAGCTGCAAGAGTGTCTGAGCAGCGCAGATTAAGCCAACT	2683
Dd	22731	CCACTGTGAATGATTAAGCTGAGCTGCAAGAGTGTCTGAGCAGCGCAGATTAAGCCAACT	22672
Oy	2684	TCACCAAGTAGAGACCATTAACACCGACGGTCAAACTCTAATAAGCAGGGCAAGACACAG	2743
Dd	22671	TCACCAAGTAGAGACCATTAACACCGACGGTCAAACTCTAATAAGCAGGGCAAGACACAG	22611
Oy	2744	ATTTTCCCCGCTTCATGAACAGAGAGACATCTCTGTGTGCTACCTGAAATGGAAGAGG	2803
Dd	22611	ATTTTCCCCGCTTCATGAATGAGAAAGAGACATCTCTGTGTGCTACCTGAAATGGAAGAG	22552
Oy	2804	TGTTTGGCTTCCCGTCCACTACAGAGCTC 2835	
Dd	22551	TGTTTGGCTTACCTGTCTACTACAAAGACC 22520	

RESULT 10
BC018214

RESULT 10
BC018214

LOCUS BC018214 1758 bp mRNA linear PRI 06-DEC-2001
 DEFINITION Homo sapiens, clone IMAGE:3862699, mRNA, partial cds.
 ACCESSION BC018214
 VERSION BC018214.1 GI:17390481
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 1758)
 AUTHORS Strausberg, R.
 JOURNAL Direct Submission
 Submitted (03-DEC-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 REMARK Contact: MGC help desk
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: DCTD/DPF
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: villalobcm.tmc.edu,
 Villalob, D.K., Luna, R.A., Hale, S.M., Huylk, S., Lu, X., Garcia,
 A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
 Muzny, D.M., Gibbs, R.A.
 FEATURES
 source
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRM Plate: 21 Row: n Column: 23.
 Location/Qualifiers
 1..1758
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3862699"
 /tissue="type="Ovary", adenocarcinoma"
 /clone_lib="NIH_MGC_66"
 /lab_host="DH10B"
 /note="Vector: pCMV-Sport6"
 <1..859
 /codon_start=2
 /product="unknown (protein for IMAGE:3862699)"
 /protein_id="AAH18214.1"
 /db_xref="GI:17390482"
 /translation="AERKRPVRLSIFDGIATGLILVLDIGIOVDRIASEVCEDSIT
 VGMVHOCKIMVGVDSVYOKHIOEMGFEDLVIGSCNLSIYNAPRGLEYEGTGR
 LFEFFYRLIHARBGDDPRPMLEMYNAGVADKRDISTISVNPATDKKPYSA
 AHKAYPNCINPGMREPLASTVNDKLEIODECEHRIKAKSKVYTTITTBNSKKOKO
 OHFVPMKEDDILMCTEMERVRGPFVHTTVDVSNRSLRORLGRSMSPVIRHLFA
 PKEYFACY"
 CDS
 BASE COUNT 513 a 376 c 428 g 441 t
 ORIGIN
 Query Match 27.8%; Score 1164.8; DB: 9; Length 1758;
 Best Local Similarity 86.2%; Pred. No. 4.9e-267;
 Matches 1483; Conservative 0; Mismatches 172; Indels 65; Gaps 15;

2265 CCGCAGCGTCACACAGAGCATATCCAGAGAGTGGGGCCCATTCGACCTGGTGAATGGAG 2324
 |||||
 Db 181 CCGCAGCGTCACACAGAGCATATCCAGAGAGTGGGGCCCATTCGACCTGGTGAATGGAG 240
 |||||
 Qy 2325 CAGTCCGTGCAATAGCTCTCCATGTCACACCTGCCCCGAGGAGCTTTATGAGGTGAC 2384
 |||||
 Db 241 CAGTCCGTGCAATAGCTCTCCATGTCACACCTGCCCCGAGGAGCTTTATGAGGTGAC 300
 |||||
 Qy 2385 TGGCGCGCTCTTTTGAATTTACCGCTCCTGCATATATGCGCGCCCAAGAGGAGAA 2444
 |||||
 Db 301 TGGCGCGCTCTTTTGAATTTACCGCTCCTGCATATATGCGCGCCCAAGAGGAGAA 360
 |||||
 Qy 2445 TGAATGCGCCCTTCTTGTGGCTCTTTGAGAAATGTGTGGCCATGGCGCTTATGACAAAG 2504
 |||||
 Db 361 TGAATGCGCCCTTCTTGTGGCTCTTTGAGAAATGTGTGGCCATGGCGCTTATGACAAAG 420
 |||||
 Qy 2505 GGACATCTCGGATTTCTTGAATCTAACCCTGTATGATGACGCCAAGAGTGTCTGC 2564
 |||||
 Db 421 GGACATCTCGGATTTCTTGAATCTAACCCTGTATGATGACGCCAAGAGTGTCTGC 480
 |||||
 Qy 2565 TGCACACAGGCGCGTACTTCTGGGGTAACTTCTGTCATGAACAGCGCTTGGCATC 2624
 |||||
 Db 481 TGCACACAGGCGCGTACTTCTGGGGTAACTTCTGTCATGAACAGCGCTTGGCATC 540
 |||||
 Qy 2625 CAGTGTGAATGATTAAGCTGGAGCTGGCAAGAGTGTGGAGGACGAGCAATATGACCAAGTT 2684
 |||||
 Db 541 CAGTGTGAATGATTAAGCTGGAGCTGGCAAGAGTGTGGAGGACGAGCAATATGACCAAGTT 600
 |||||
 Qy 2685 CAGCAATAGAGAGCAATTAACACAGCTGCAACCAATCTAAGAGAGGAGCAAGAGCAAGA 2744
 |||||
 Db 601 CAGCAATAGAGAGCAATTAACACAGCTGCAACCAATCTAAGAGAGGAGCAAGAGCAAGA 660
 |||||
 Qy 2745 TTTCCCGCTCTTCAATGAACAGAGAGCAATCTGTGTGTCACCTGAATGAAAGGCT 2804
 |||||
 Db 661 TTTCCCGCTCTTCAATGAATGAAGAGAGCAATCTGTGTGTCACCTGAATGAAAGGCT 720
 |||||
 Qy 2805 GTTGGCTCCCGCTCCATGACATGACAGAGCTGTCAACATGAGCGCTTGGCGAGGAGAG 2864
 |||||
 Db 721 ATTTGGTTTCCCATGTCACATATGACATGTCACCAATGAGCGCTTGGCGAGGAGAG 780
 |||||
 Qy 2865 ACTGTGGGCGCATGTCGAGAGCTGCGGTCATCCGCCACCTCTTCCCTCGCTGAAGAA 2924
 |||||
 Db 781 ACTGTGGGCGCATGTCGAGAGCTGCGGTCATCCGCCACCTCTTCCCTCGCTGAAGAA 840
 |||||
 Qy 2925 ATATTTTCTGTGTGTAAGAGGACATGGGAGCAAACTGAAGTGAATGATTAAGAAATT 2984
 |||||
 Db 841 ATATTTTCTGTGTGTAAGAGGACATGGGAGCAAACTGAAGTGAATGATTAAGAAATT 896
 |||||
 Qy 2985 AAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 3044
 |||||
 Db 897 AAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 951
 |||||
 Qy 3045 AGAAGAGT-TGAGCAGCCAG 3101
 |||||
 Db 952 AGAAGAGTATCAGACCCAG 1011
 |||||
 Qy 3102 AAACCGCGAGAGCGCTGCGCTTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3161
 |||||
 Db 1012 AAACCGCGAGAGCGCTGCGCTTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1071
 |||||
 Qy 3162 TAACCTTGAATCTATTAAG 3217
 |||||
 Db 1072 TAACCTTGAATCTATTAAG 1131
 |||||
 Qy 3218 TCCATGAG 3277
 |||||
 Db 1132 TCCATGAG 1190
 |||||
 Qy 3278 GTTTTGTCTCTGCTGAG 3336
 |||||
 Db 1191 G-TTTTGTCTCTGCTGAG 1249
 |||||
 Qy 3337 AAAAAGTAAAGATGCTTGGAG 3387
 |||||

Db	1250	AAATAGTACATACCTTGCAGAGAAAGCTGGAGAGAGAAAAAGAAATCTATAG	1309
QY	3388	AAACTTAAATATGTGGTTTTTTTTTTTTTTCCTTT	3422
Db	1310	AAATCTATATATGGTGTGTTTTTTTTTTTGTATTTTTTTTGGGTTTTTTT	1369
QY	3423	-----TCTATATATCTCTTTGGTTGTCTAGCCTATCAGATGAGACCAAAACA--G	3474
Db	1370	TTTACTATATATCTTTTTTTTTTGTGCTCTAGCCTATCAGATGAGACCAAGCAGG	1429
QY	3475	GAAGAAATGAGACCCCTGGAGGCAG--AGCTCCTCTCCACCCCGGAGCACTGCAA	3533
Db	1430	GACGGAAAGAGAGACACTCAGCGGACAGCATTCCTCCAGCACTGAGCTGCTGC	1489
QY	3534	CAGCACCATTCTGTGTCATGCAAAAACAACCACTATGACGACGAGCGCTGAGAGAA	3593
Db	1490	CAGACCATTCCTGTGTGTCACGCCAAAACAAGAACCCGTTATGCGACGAGGAGACAGAACACC	1549
QY	3594	CCACACACGACACTTCTTACAGTATTTTTCAGGTGCTTACACACAGAAACCTTGAAGAA	3653
Db	1550	ACACAAAGACA-TTTTTCTACAGTATTTCCAGGTGCTTACACACAGAAACCTTGAAGAA	1608
QY	3654	ACCACTTTCTAGAAGCCGCTGTATCTCTCTGTGTTACAGTTTATATATATGATAGTAT	3713
Db	1609	ATCAGTTTCTAGAAGCCGCTGTATCTCTCTGTGTTACAGTTTATATATATGATAGTAT	1668
QY	3714	GAGATATATATATATTAAGTACTGTAACTACTGTACA	3753
Db	1669	GAG--ATATATATATTAAGGCTACTGTAACTACTGTAAA	1706

RESULT	11
ACI03454/c	
LOCUS	
DEFINITION	ACI03454 bp DNA linear HTG 12-JUL-2002
ACCESSION	ACI03454
VERSION	ACI03454.3 GI:21728739
KEYWORDS	HTGS; PHASEL.
SOURCE	Norway rat.
ORGANISM	Rattus norvegicus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus. 1. (bases 1 to 117264) Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alisbrooks,S.L., Amaraltinge,H.C., Ate,J.T.R., Ayale,M., Banks,T., Bardella,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bonick,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Butley,C., Burck,P., Burkett,C., Burrill,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dabhorne,S.R., David,R., Dayala,M.L., Davis,C., Davy-carroll,L., Deedrich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earhart,C., Edgar,D., Edwards,C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorell,U.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues-Mc, Hollaway,C., Hollins,B., Homsl,F., Howard,S., Huber,J., Huliy,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Joliviet,S., Jouah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratochvic,J., Kunesht,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Litcharge,O., Lieu,C., Liu,J., Liu,W., Louisseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Matheswarai,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhinney,E., Mcloed,M.P., Meador,M., Mel,G., Metzke,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,A., Morris,S., Mosser,M., Neel,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,

Nguyen, N., Nickerson, E., Nwokkenko, S., Ogunb, M., Okunnu, G.,
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quties, M., Ren, Y.,
Rivers, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G.,
Schärer, S., Scott, G., Shen, H., Shoochster, N., Sisson, I.,
Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Swatek, A., Tabors, P., Tameris, A., Tameris, K., Tang, H.,
Tansley, J., Taylor, C., Taylor, T., Teliford, B., Thomas, R., Thomas, S.,
Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, N., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleciyk, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 117264)
Worley, K.C.

Direct Submission
Submitted (25-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 117264)
Worley, K.C.

Direct Submission
Submitted (12-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 11, 2002 this sequence version replaced gi:17973284.

----- Genome Center -----
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information -----
Center project name: GHPN
Center clone name: CH230-130117
----- Summary Statistics -----
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 54254 bases at least Q40
Consensus quality: 56889 bases at least Q30
Consensus quality: 59339 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_drift_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 60 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

*	1	1343:	contig of 1343 bp in length
*	1344	1443:	gap of unknown length
*	1444	2582:	contig of 1139 bp in length
*	2583	2682:	gap of unknown length
*	2683	3717:	contig of 1035 bp in length
*	3718	3817:	gap of unknown length
*	3818	4912:	contig of 1095 bp in length
*	4913	5012:	gap of unknown length
*	5013	6349:	contig of 1337 bp in length
*	6350	6449:	gap of unknown length
*	6450	7874:	contig of 1425 bp in length
*	7875	7974:	gap of unknown length
*	7975	9050:	contig of 1076 bp in length
*	9051	9150:	gap of unknown length
*	9151	10615:	contig of 1465 bp in length
*	10616	10715:	gap of unknown length
*	10716	12258:	contig of 1543 bp in length
*	12259	12358:	gap of unknown length
*	12359	13675:	contig of 1317 bp in length

13676 13775: gap of unknown length
13776 15423: contig of 1648 bp in length
15424 15523: gap of unknown length
15524 16891: contig of 1368 bp in length
16892 16991: gap of unknown length
16992 18433: contig of 1442 bp in length
18434 18533: gap of unknown length
18534 18559: contig of 1326 bp in length
18560 19959: gap of unknown length
19960 21010: contig of 1051 bp in length
21011 21110: gap of unknown length
21111 22710: contig of 1600 bp in length
22711 22810: gap of unknown length
22811 24047: contig of 1237 bp in length
24048 24147: gap of unknown length
24148 25330: contig of 1183 bp in length
25331 26558: contig of 1128 bp in length
26559 26659: gap of unknown length
26660 28182: contig of 1524 bp in length
28183 28282: gap of unknown length
28283 29368: contig of 1086 bp in length
29369 29468: gap of unknown length
29469 30746: contig of 1278 bp in length
30747 30846: gap of unknown length
30847 32194: contig of 1348 bp in length
32195 32294: gap of unknown length
32295 33410: contig of 1116 bp in length
33411 33510: gap of unknown length
33511 33612: contig of 1502 bp in length
33613 33712: gap of unknown length
33713 36244: contig of 1132 bp in length
36245 36344: gap of unknown length
36345 37377: contig of 1033 bp in length
37378 37477: gap of unknown length
37479 39062: contig of 1585 bp in length
39063 39162: gap of unknown length
39163 40659: contig of 1497 bp in length
40660 40759: gap of unknown length
40760 41926: contig of 1167 bp in length
41927 42026: gap of unknown length
42027 43382: contig of 1256 bp in length
43383 43987: gap of unknown length
43988 44987: contig of 1605 bp in length
44989 45087: gap of unknown length
45088 46483: contig of 1396 bp in length
46484 47680: gap of unknown length
47681 47780: gap of unknown length
47782 49170: contig of 1350 bp in length
49171 49270: gap of unknown length
49272 51444: contig of 2174 bp in length
51445 51544: gap of unknown length
51546 52836: contig of 1252 bp in length
52837 52936: gap of unknown length
52937 55045: contig of 2109 bp in length
55046 55145: gap of unknown length
55146 56691: contig of 1546 bp in length
56692 56791: gap of unknown length
56792 58443: contig of 1652 bp in length
58444 58543: gap of unknown length
58545 60538: contig of 1995 bp in length
60539 60638: gap of unknown length
60640 63571: contig of 2933 bp in length
63572 63671: gap of unknown length
63673 65119: contig of 1448 bp in length
65120 65219: gap of unknown length
65220 66996: contig of 1777 bp in length
66997 67096: gap of unknown length
67097 69660: contig of 2564 bp in length
69661 72469: gap of unknown length
72470 72569: contig of 2709 bp in length

Query Match 25.2% Score 1055.2; DB 2; Length 117264;
Best Local Similarity 90.4% Pred. No. 1.1e-240;
Matches 1304; Conservative 0; Mismatches 70; Indels 68; Gaps 14;
* 72570 75117: contig of 2548 bp in length
* 75118 75217: gap of unknown length
* 75218 76930: contig of 1713 bp in length
* 76931 77030: gap of unknown length
* 77031 78952: contig of 1922 bp in length
* 78953 79052: gap of unknown length
* 79053 82025: contig of 2973 bp in length
* 82026 82125: gap of unknown length
* 82126 83473: contig of 1348 bp in length
* 83474 83573: gap of unknown length
* 83574 86343: contig of 2770 bp in length
* 86344 86444: gap of unknown length
* 86444 90103: contig of 3660 bp in length
* 90104 90203: gap of unknown length
* 90204 95081: contig of 4878 bp in length
OY 2798 AAGGGTGTGGCTCCCGCTCAGTACAGAGCGTCCACATGAGCGCGTGGCGA 2857
DB 105266 AAGGGTGTGGCTCCCGCTCAGTACAGAGCGTCCACATGAGCGCGTGGCGA 105207
OY 2858 GGCAGAGCTGCTGGCGGCGATGCGAGCGCGGTCATCCGCGACCTTCGCGCG 2917
DB 105206 GGCAGAGCTGCTGGCGGCGATGCGAGCGCGGTCATCCGCGACCTTCGCGCG 105147
OY 2918 TGAAGAAATTTTGTCTGTGTGTAAAGGACATGGCGCAACTGAATGATGATTA 2977
DB 105146 TGAAGAAATTTTGTCTGTGTGTAAAGGACATGGCGCAACTGAATGATGATTA 105087
OY 2978 TGAAGAAATTTTGTCTGTGTGTAAAGGACATGGCGCAACTGAATGATGATTA 2977
DB 105086 CCAAGTTAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 105037
OY 3038 AGGACGAGGAAAGATTCAGACCCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 3097
DB 105036 AGGACGAGGAAAGATTCAGACCCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 104977
OY 3098 AGGAAAGCGCGGAGGCTTGGCTTGAAGAGGTTGACATCATCTCCTGATTTTCA 3157
DB 104976 AGG-AAATGCGGAGGCTTGGCTTGAAGAGGTTGACATCATCTCCTGATTTTCA 104918
OY 3158 ATGTTACCTTCACTCTATCTAAAGCAAAATAGGCGCCCTCCCTCCCTCCGCG 3217
DB 104917 ATGTTACCTTCACTCTATCTAAAGCAAAATAGGCGCCCTCCCTCCCTCCGCG 104858
OY 3218 TCCATGAGGCGGAACTTTTGTCTTCTACTC--TTTTCAGAGGGGTTTGTGTTT 3275
DB 104857 TCCATGAGGCGGAACTTTTGTCTTCTACTC--TTTTCAGAGGGGTTTGTGTTT 104798
OY 3276 GGGTTTTTGTCTGTCTGTGACTGAACAGAGAGTT-ATTGACGAAATTCAGTAACA 3334
DB 104797 GGGTTTTTGTCTGTCTGTGACTGAACAGAGAGTT-ATTGACGAAATTCAGTAACA 104738
OY 3335 ACAGAAAGTAAATGCTTGGAGAGGAAAGGAGAGGAGGAAATTCATTAATAACTTA 3394
DB 104737 ACAGAAAGTAAATGCTTGGAGAGGAAAGGAGAGGAGGAAATTCATTAATAACTTA 104681
OY 3395 AAATATGTTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTT 3454
DB 104680 AAATATGTTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTT 104624
OY 3455 TCAGATGAGGAGCAACAGAGGAAATAGAGAGCCCTGGAGGAGAGTCTCTCTCC 3514
DB 104623 TCAGATGAGGAGCAACAGAGGAAATAGAGAGCCCTGGAGGAGAGTCTCTCTCC 104564
OY 3515 ACCGCGGAGAGTCTACAGACATTCCTGGTCAATGCAAAAGCAAGCAAGCAAGCAAGCA 3574
DB 104563 ACCGCGGAGAGTCTACAGACATTCCTGGTCAATGCAAAAGCAAGCAAGCAAGCAAGCA 104504
OY 3575 CGAGGGGCTGAGAGCAACACACACAGCACTTTTACAGATTTTTCAGGTGCTTACAC 3634

```

Db      104503 GCAGGGCGCTGAGATACACACAGACAC-TTTTACAGTATTTGAGGCTTACCAC 104445
Qy      3635 ACAGGAAACCTTGAGAAACCAAGCTTCTTGAGAGCCCTGTACCTTTGTTACAGTTT 3634
Db      104444 ACAGGAAACCTTGAGAAACCAAGCTTCTTGAGAGCCCTGTACCTTTGTTACAGTTT 104385
Qy      3695 ATATATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3754
Db      104384 ATATATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 104327
Qy      3755 CCCGACTTCAATATGCTGCTTTCAAAACAGAGATGAGCAAAACATCAGCTTCCGCT 3814
Db      104326 CCCAACTTCAATATGCTGCTTTCAAAACAGAGATGAGCAAAACATCAGCTTCCGCT 104267
Qy      3815 GGGCCCTGTGCAAAAGGTTTCAAGCCAGATGGGAGAGGGAGCGCTGAGAG----- 3859
Db      104266 GGGCCCTGTGCAAAAGGTTTCAAGCCAGATGGGAGAGGGAGCGCTGAGAG----- 104208
Qy      3870 -----GGGTTTAACTGAGAGATGAGCCATAT- 3900
Db      104207 GTTTTTTTTTTTTTTTTTTGTGCTGCTGTTTAACTGAGAGATGAGCCATAT- 104148
Qy      3901 ---CACCCCCACCCCTGCCCCATGCTTACCTGCTGCAAAA-AGGGGCTGAGCTGA 3956
Db      104147 GCGCCACACCCACCCACCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 104088
Qy      3957 GGTGCTGAGACCCCTGAGGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 4016
Db      104087 GGTGCTGAGACCCCTGAGGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 104028
Qy      4017 TAGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4071
Db      104027 TTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 103968
Qy      4072 CCTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4130
Db      103967 CCTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 103908
Qy      4131 AGTAGGAGAAAGTACCCGCTATGATGATGATGATGATGATGATGATGATGATGATGAT 4190
Db      103907 AGTAGGAGAAAGTACCCGCTATGATGATGATGATGATGATGATGATGATGATGATGAT 103848
Qy      4191 GT 4192
Db      103847 GT 103846

```

RESULT 12
AC009474 168651 bp DNA linear PRI 08-NOV-2000
LOCUS Homo sapiens BAC clone RP11-179G23 from 2, complete sequence.
ACCESSION AC009474 GI:8748849
VERSION AC009474.4
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 168651)
AUTHORS Sulston, J.E. and Waterston, R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
PUBMED 9847074
REFERENCE 2 (bases 1 to 168651)
AUTHORS Mulvaney, E., Stoneking, T., Lapiant, Y. and McDill, B.
TITLE The sequence of Homo sapiens BAC clone RP11-179G23
JOURNAL Unpublished
AUTHORS 3 (bases 1 to 168651)
Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (24-AUG-1999) Genome Sequencing Center, Washington

University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE 4 (bases 1 to 168651)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (27-JUN-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE 5 (bases 1 to 168651)
AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (08-NOV-2000) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Jun 27, 2000 this sequence version replaced gi:7630880.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: saplens@wustl.wustl.edu
----- Summary Statistics
Center project name: H_NH0179G23

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality > 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
The RP11-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Frenken, E., Tatenno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)
VECTOR: pBAC3.6
NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-547F18; the clone sequenced to the right is RP11-444B4, 200 base pair overlap. Actual end of this clone is at base position 18711 of RP11-444B4.

FEATURES
Location/Qualifiers
1..168651
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-179G23"
/clone_lib="RP11-11"
13..497
/rpl_family="L2"
616..683
/rpl_family="L2"
1615..1815
/rpl_family="MIR"
2577..2748
/rpl_family="MIR"
2954..3144
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region

RESULT 13
 AC112040/c
 LOCUS
 DEFINITION
 AC112040
 *** 59 unordered pieces.
 AC112040.2 GI:21737402
 HTG: HTGS_PHASE1.
 KEYWORDS
 SOURCE
 ORGANISM
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE

AUTHORS

1 (bases 1 to 119630) DNA linear HTG 13-JUL-2002
 Muzny,D.M., Adams,C., Adio-Oduola,B., Alt-osman,F.R., Allen,C.,
 Aisbrooks,S.L., Amaralunga,H.C., Are,J.R., Avela,M., Banks,T.,
 Barberis,J., Benton,J., Blumhage,K., Blankenburg,K., Bonnin,D.,
 Bouck,J., Bowie,S., Brilews,M., Brown,E., Brown,M., Bryant,N.P.,
 Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
 Carron,T.F., Carter,M., Cavazos,S.R., Chacko,U., Chavez,D.,
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
 Cleveland,C.D., Cox,C., Coyle,M.D., Dalhorne,S.R., David,R.,
 Devilla,M.L., Davis,C., Davy-Carroll,L., Deckerich,D.A.,
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
 Douthett,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
 Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
 Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
 Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
 Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
 Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,
 Homaj,F., Howard,S., Huber,J., Huylk,S., Hume,J., Jackson,L.E.,
 Jacobson,B., Jia,Y., Johnson,R., Jolyet,S., Joudah,S.,
 Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
 Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
 Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
 Lozador,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
 Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
 Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Metzger,M.,
 Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
 Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
 Nguyen,N., Nickerson,E., Nwokkwo,S., Ogub,M., Okunnu,G.,
 Otaguave,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
 Peters,L., Plickens,R., Primus,E., Pull,L., Quiles,M., Ren,Y.,
 Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,
 Scherer,S., Scott,G., Shen,H., Shooshair,N., Sisson,I.,
 Sodergren,E., Soudake,T., Sparks,A., Stanley,H., Stone,H.,
 Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
 Tansey,J., Taylor,C., Taylor,T., Telifod,B., Thomas,N., Thomas,S.,
 Usmani,K., Vasquez,L., Vera,V., Villalob,D., Vinson,R., Wang,Q.,
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,
 Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
 Weinstein,G. and Gibbs,R.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE
 TITLE
 JOURNAL
 AUTHOR
 COMMENT
 1 (bases 1 to 119630)
 Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 119630)
 Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jul 12, 2002 this sequence version replaced gi:18701992.
 ----- Genome Center
 Center: Baylor College of Medicine

Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 Project Information
 Center project name: GPR0
 Center clone name: CH230-61E10
 Summary Statistics
 Sequencing vector: Plasmid:
 Chemistry: Dye-terminator Big Dye: 100% of reads
 Assembly program: Phrap: version 0.990329
 Consensus quality: 48405 bases at least Q40
 Consensus quality: 52337 bases at least Q30
 Consensus quality: 56201 bases at least Q20

 NOTE: Estimated insert size may differ from sequence length
 (see <http://www.hgsc.bcm.tmc.edu/docs/genbank/draft.html>).
 NOTE: This is a 'working draft' sequence. It currently
 consists of 59 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

1	1359:	contig of 1359 bp in length
1360	1459:	gap of unknown length
1460	2850:	contig of 1391 bp in length
2851	2950:	gap of unknown length
2951	4544:	contig of 1594 bp in length
4545	4645:	gap of unknown length
4645	5975:	contig of 1331 bp in length
5976	6075:	gap of unknown length
6076	7435:	contig of 1360 bp in length
7436	8535:	gap of unknown length
8536	8635:	gap of unknown length
8636	10194:	contig of 1559 bp in length
10195	10294:	gap of unknown length
10295	11310:	contig of 1016 bp in length
11311	11410:	gap of unknown length
11411	12651:	contig of 1241 bp in length
12652	12751:	gap of unknown length
12752	14281:	contig of 1530 bp in length
14282	14381:	gap of unknown length
14382	15543:	contig of 1162 bp in length
15544	16643:	gap of unknown length
16644	17091:	contig of 1448 bp in length
17092	17191:	gap of unknown length
17192	18716:	contig of 1525 bp in length
18717	18816:	gap of unknown length
18817	20199:	contig of 1363 bp in length
20200	21839:	gap of unknown length
21840	21939:	contig of 1540 bp in length
21940	23412:	gap of unknown length
23413	23512:	contig of 1473 bp in length
23513	24818:	gap of unknown length
24819	24918:	contig of 1306 bp in length
24919	26138:	gap of unknown length
26139	26238:	contig of 1220 bp in length
26239	27593:	gap of unknown length
27594	27693:	contig of 1355 bp in length
27695	29552:	gap of unknown length
29553	29652:	contig of 1859 bp in length
29654	31301:	gap of unknown length
31302	31401:	contig of 1649 bp in length
31402	32850:	gap of unknown length
32851	32950:	contig of 1458 bp in length
32951	34268:	gap of unknown length
34269	34369:	contig of 1310 bp in length
34370	35833:	gap of unknown length
35834	35933:	contig of 1464 bp in length
35934	37227:	contig of 1294 bp in length

Query Match	Best Local Similarity	Score	DB 2;	Length	119630;
Matches 1108;	Conservative	0;	Mismatches 60;	Indels 30;	Gaps 8;
37328	37327:	gap of	unknown	length	
37328	38672:	config	of 1345 bp	in length	
38673	38772:	gap of	unknown	length	
38773	40409:	config	of 1637 bp	in length	
40410	40509:	gap of	unknown	length	
40510	42386:	config	of 1877 bp	in length	
42387	42486:	gap of	unknown	length	
42487	43891:	config	of 1405 bp	in length	
43892	43991:	gap of	unknown	length	
43992	45547:	config	of 1556 bp	in length	
45548	45647:	gap of	unknown	length	
45648	47802:	config	of 2155 bp	in length	
47803	47902:	gap of	unknown	length	
47903	49367:	config	of 1465 bp	in length	
49368	49467:	gap of	unknown	length	
49468	51732:	config	of 2265 bp	in length	
51733	51832:	gap of	unknown	length	
51833	53464:	config	of 1633 bp	in length	
53465	53564:	gap of	unknown	length	
53565	55043:	config	of 1479 bp	in length	
55044	55143:	gap of	unknown	length	
55144	56671:	config	of 1528 bp	in length	
56672	56771:	gap of	unknown	length	
56772	58608:	config	of 1837 bp	in length	
58609	58708:	gap of	unknown	length	
58709	60082:	config	of 1374 bp	in length	
60083	60182:	gap of	unknown	length	
60183	61662:	config	of 1480 bp	in length	
61663	61762:	gap of	unknown	length	
61763	64208:	config	of 2446 bp	in length	
64209	64308:	gap of	unknown	length	
64309	66380:	config	of 2072 bp	in length	
66381	66480:	gap of	unknown	length	
66481	66078:	config	of 2598 bp	in length	
66079	69178:	gap of	unknown	length	
69179	71272:	config	of 2094 bp	in length	
71273	71372:	gap of	unknown	length	
71373	73377:	config	of 2005 bp	in length	
73378	73477:	gap of	unknown	length	
73478	75898:	config	of 2421 bp	in length	
75899	75998:	gap of	unknown	length	
75999	77816:	config	of 1818 bp	in length	
77817	77916:	gap of	unknown	length	
77917	80747:	config	of 2831 bp	in length	
80748	80847:	gap of	unknown	length	
80848	82669:	config	of 1822 bp	in length	
82670	82769:	gap of	unknown	length	
82770	84419:	config	of 1650 bp	in length	
84420	84519:	gap of	unknown	length	
84520	87705:	config	of 3186 bp	in length	
87706	87805:	gap of	unknown	length	
87806	90005:	config	of 2200 bp	in length	
90006	90105:	gap of	unknown	length	
90106	92538:	config	of 2433 bp	in length	
92539	92638:	gap of	unknown	length	
92639	95113:	config	of 2475 bp	in length	
95114	95213:	gap of	unknown	length	
95214	98154:	config	of 2941 bp	in length	

Db	56206	CCTGGCATGAACAGCCCACTTGGCAATTACACTGTGAATGATATAGCTGGAGCTGCCAAGAGTGT	5614
QY	2659	CTGGAGCAGCGCAGCAATATAGCCAACTTTCAGCAAACTGAGGACCATTACACCAGGTCAAC	2718
Db	56146	CTGGAAACACG-----GCCAACTTCAGCAAACTGAGGACCATTACACCAGGTCAAC	5609
QY	2719	TCCTATTAAGCAGGCGAAGACACAGCACTTTCGCCGTCTTCATGAACGAGAAAGGAGACATC	2778
Db	56094	TCCATTAAGCAGGCGCAAGACACAGCACTTTCGCCGTCTTCATGAATGAGAAAGAGACATC	5603
QY	2779	CTGGAGTGCACATGAAATGGAAGGGGTGTTGGCTTCCCGCCCACTACAGAGCTGCTC	2838
Db	56034	CTGTGGTGCACATGAAATGGAAGGGGTGTTGGCTTCCCGCCCACTACAGAGCTGCTC	5597
QY	2839	AACATGAGCCCTTGGGAGGACAGACAGCTGTGGGCGCATGCTGGAGACCTGCGCGTCATC	2898
Db	55974	AACATGAGCCCTTGGGAGGACAGACAGCTGTGGGCGCATGCTGGAGACCTGCGCGTCATC	5591
QY	2899	CGCCACCTCTTCCGCTCCGCTGAAGGAATATTTTGCTGTGTGAAGGACATGGGGCA	2958
Db	55914	CGCAACCTCTTCCGCGCCGCTGAAGGAATATTTTGCTGTGTGAAGGACATGGGGCA	5585
QY	2959	ACTGACAGTATGATATATTAATACTTTAAACAACAACAACAACAACAACAACAACA	3018
Db	55854	ACTGAGATGATGATATATTAACCAACTTTAAACAACAACAACAACAACA-----ACACA	5580
QY	3019	ATTAACACCAAGAACGAGAGGAGCGGAGAAAGTTCCAGACCCCGAAGAGAGAAAGGAT	3078
Db	55804	ATTAACACCAAGAACGAGAGGAGCGGAGAAAGTTCCAGACCCCGAAGAGAGAAAGGAT	5574
QY	3079	TTAAAGCAACACAGAGAGGAGAAAGCGCCGAGGGCTTGCCCTTGCAAAAAGGTTGGAC	3138
Db	55744	TTAATGCAAAACACAGAGAGAGG-AAAAGCCGAGGGCTTGCCCTTGCAAAAAGGTTGGAC	5568
QY	3139	ATCATCTCCCTAGATTTCAATGTTAACTTCAGTCCATCTCACTAAAGCAAAATGGGCC	3198
Db	55685	ATCATCTCTAGATTTCAATGTTCACTTCATCTCACTAAAGCAAAAGGATGGGCC	5562
QY	3199	TCCCTCTTCCCTCCGCGTCTTGAAGAGGCAACTTTTGTCTTCACTCTTTTTCAGAG	3258
Db	55625	TCCCTCTTCCCTCCGCGTCTTGAAGAGGCAACTTTTGTCTTCACTCTTTTTCAGAG	5556
QY	3259	-GGCTTTTCTGTGTGTGGGTTTTTGTCTGTCTGTGACTGAAACAGAGATT-ATT	3316
Db	55565	AGGGGTTCTGTGTGTGGGTTTTTGTCTGTCTGTGACTGAAACAGAGATT-ATT	5550
QY	3317	GCAGCAAAATCAGTATACACAAAAAGTAAATGCTTTGGAGAGAGAAAGGAGAGAGGA	3376
Db	55505	GCAGCAAAATCAGTATACACAAAAAGTAAATGCTTTGGAGAGAGAAAGGAGAGAGGA	5544
QY	3377	AAATCTCTTAAAACTTAAATATTTGGTTTTTTTTTTTTTCTTCTATATATCTCTT	3436
Db	55445	AAATCTCTTAAAACTTAAATATTTG-----TTTTTTTTTCTTCTATATATCTCTT	5539
QY	3437	TGTTGTCTCAGCTGTATAGAGACCAAAACAGAGAGAAATAGAGACCCTCGGA	3496
Db	55391	TGTTGTCTCAGCTGTATAGAGACCAAAACAGAGAGAAATAGAGACCCTCGGA	5533
QY	3497	GCGAGAGTCTCTCTCCACCCCGCAGACAGTCTCAACAGACCACTTCGTGGTATGCA	3556
Db	55331	GCGAGAGTCTCTCTCTCCACCCCGCAGACAGTCTCAACAGACCACTTCGTGGTATGCA	5527
QY	3557	AACAGAACCAACTAGCAGCAGAGGGGCGTGAAGAAACACACACAGACAC-TTTTCTACAG	3615
Db	55271	AACAGAACCAACTAGCAGCAGAGGGCGT--GAGAAACACACACACAGACACTTTTCTACAG	5521
QY	3616	TATTTACAGTGCCTACACACAGGAAACCTTGAAGAAAAACAGTTTCTAGAAAGCGCT	3673
Db	55213	TATTTACAGTGCCTACACACAGGAAACCTTGAAGAAAAACAGTTTCTAGAAAGCGCT	5516

LOCUS	AF068627	4135 bp	mRNA	linear	ROD 06-DDC-1999
DEFINITION	Mus musculus DNA cytosine-5 methyltransferase 3B2 (Dmnc3b) mRNA,				
ACCESSION	AF068627				
KEYWORDS	alternatively spliced, complete cds.				
SOURCE	AF068627.2 GI:6449471				
ORGANISM	Mus musculus.				
AUTHORS	Eukariota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
TITLE	1 (bases 1 to 4135) Okano, M., Xie, S., and Li, E. Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases				
JOURNAL	Nat. Genet. 19 (3), 219-220 (1998)				
REFERENCE	9632476 2 (bases 1 to 4135) Xie, S., Okano, M., and Li, E. Direct Submission Submitted (28-MAY-1998) CVRC, Mass. Gen. Hospital, 149 13th Street, Charlestown, MA 02129, USA 3 (bases 1 to 4135) Okano, M., Chijiwa, T., Sasaki, H., and Li, E. Direct Submission Submitted (04-NOV-1999) CVRC, Mass. Gen. Hospital, 149 13th Street, Charlestown, MA 02129, USA Sequence update by submitter On Nov 18, 1999 this sequence version replaced gi:3327981.				
REMARK	COMMENT FEATURES source				
gene	1. .4135 /organism="Mus musculus" /db_xref="taxon:10090" /map="84.0 cm" 1. .4135 /gene="Dmnc3b"				
misc_feature	1. .4135 /note="Dmnc3b" /note="similar to EST sequences deposited in GenBank Accession Numbers A116594, A119197, A117277, A120568, A407106, and A457617" 269. .2788 /gene="Dmnc3b" /function="de novo DNA methylation" /note="alternatively spliced product; contains Cys-rich region; C-terminal region is similar to corresponding region of DNA cytosine-5 methyltransferase 3A" /codon_start=1 /product="DNA cytosine-5 methyltransferase 3B2" /protein_id="AA40179.2" /db_xref="GI:6449472"				
CDS	1/translation="MKDSRLHNEEGAGSYDECIIVNCFSDSDTRKADSPVLE LIEETEDTRRESEAVATRHNSGTSLEKREVSILANTQDMGDDRDYDONGSDILMPK SRRRRSSASSTPSPASVDMEETVTPSTVSDSDSDGDMTQVDAESD GSTEVDODKEFGIGDLVWGKIGFSPWPAMVSMKATSRROMAPGMVOMFGDFG SLSIAKILVLAIGLFSOHNLATFNKIVSYRAMVHTLEKRVACQTFSSSPGSLSD OLKDEMEHAGSEFPKPTGIEGLKPKKKOPKNSRRRTNDSASPEPRRLKSTSYSG KDGDEEERSEPMKASEVTNNKKGLKEDRCISGCKNSPSPIPLREGGLOKSDPFLPL FYWYDDGYOATCYVCEBEILLCSNTSCGCGVCLIVLIGAGTAADAKDPEPS CVYICRCHGVAPRRKDDMMRLDPPFTDPLDEPPPLVYPAIAPKRPVIVLE FDIATNGVYLKEGIVETKIVTASPECAISLVGVVAKHEQILTVMDARKITKRIEEL MGFDVLIGGSPNDLSVMPARKGLTETGRLFEETILHLNTPKRBDNPFPMW ENVAKAVAKDKDISRELACPMIDALIKVSAHAKRTFNGNLKGNRPVMSKNDKL EALDDCLSRKATLAKVOTITTKSNSITRGKNQLEPVAMGKQDVLMCTELERILGFP AHTDCLSNMGKSGVPLILGRSWSVPIRLPAPLKDYFACE"				
BASE COUNT	1043 a 1048 c 1106 g 938 t				
ORIGIN					
Query Match	19.0%; Score 796.2; DB 10; Length 4135;				
Best Local Similarity	64.8%; P: 0.64e-179;				
Matches 1267; Conservative	0; Mismatches 643; Indels 45; Gaps 4;				

OY	991	GAGCGACTAGAGGGGATGCTGGGGACAAACAAATCTACCAAAAGCAGCCGACATGAGCT	1050
Db	875	GATCAGGAGGCTATGGATATCCACACAGCTGCAATCCGAGAGACAGATGGAGACAGCA	934
OY	1051	GAGATATGAGATAGCGCGGGGCTTTGGATGGAGAGAGCTGTGGGGAAACCTGGGGG	1110
Db	935	GAGATATGAGATATTAAGAACTTTGGATATGAGCACTCGTGTGGGAAAGATCAAGGCG	994
OY	1111	TTTCTCTGCTGGGCCAGCCCAATTGTCTCTTGGTGGATGACAGGCCCGAGCCGACACT	1170
Db	995	TTTCTCTGCTGGGCCCTGCAAGTGTGTCTCTGGAAAGCCACTTCAAGCAGCAGGCAATG	1054
OY	1171	GAGGCACTCGCTGGGTCAATGTGTGTCGAGATGGCAAGTTCTCAGTGTGTGTGGAG	1230
Db	1055	CCCGAAATGGCGTGGGGTACAGTGGTGTGGTGGAGTGGCAAGTTTCTGTGAGATCTCTGTAC	1114
OY	1231	AAGCTCATGCGCGCTGACGCTCCCTTCGACGTGATTCACACAGGCGACCTTCAACAAGAG	1290
Db	1115	AAATCGTGGCTGTGGGGCTGTACGCCACACACTTATCTGGCTACCTTCAATTAAGCTG	1174
OY	1291	CCCATGTACCGCAAGACCATCATACGAAGTCTTCAGAGTGGCGACACCGCTCGCGGAAG	1350
Db	1175	GTTTCTTATATGAAAGGCCATGTACCACTCTGAGAAAGCCAGGTTGTGAGCTGGCAAG	1234
OY	1351	CTGTCTTCCACTTGGCCATGACAGTGAATGAAAGTGAACATGTGGCAAGCTGTGAGACTGAG	1410
Db	1235	ACCTTCTC-----CAGCAGTCTCGAGAGTCACTGGACACAG	1273
OY	1411	AACACAGATGATTTGAATGGGCCCTCGGTGGCTTCACGCCCTCGGCTCCTAAGGCGCTG	1470
Db	1274	CTGAAGGCCATGTGGAGTGGGCCCCACGAGGGCTTCAGACCTCATGTGGATCGAGGGCTC	1333
OY	1471	GAGCCACCAAGAAAGAAAGAAATCCTTACNAGGAATTTACACCCGACATGTGGGTGAG	1530
Db	1334	AAACCC-----CAACAAAGAACCAACACAGAAACAAATCTCMAACGCCACAAACATATAC	1387
OY	1531	CCTAAACAGACTCTTTAGGCCCCACCCGCCACAGCCAGCAAGAACCCAGAAAGACACACA	1590
Db	1388	TCTCTCTCTCTGTAGTCCGCCCCACCCACCCGCTCTCAGAACTATACATGTATGGCGGAAG	1447
OY	1591	GAGAAACCTTAAGGTCAAAGAGATCATTTGATGAGGCCACAAAGGAGGCGGTGGTATATAG	1650
Db	1448	GACCGAGGAGGATATGAGAGAGAGC-----CGAGAAAGGATGGCTTCTGAA	1492
OY	1651	GTGGGCCCAAGATGCAAAACATCATGAGAGACATTTGATCATCATGTGGAGACCTCAATATGC	1710
Db	1493	GTCAACCAACAAAGGCAATCTGGAAGACCGCTGTGTCTCTGTGAAGAAAGAACCTT	1552
OY	1711	ACCGTGAAGCACCATCTCTTCATTTGAGGCAATGTGCCAAGACTGAAGACTGCTTTCTTG	1770
Db	1553	GTGTCTCTTCCACCCCTCTTTGAAGGGGGGCTCTGTCAAGATTTGCCGGAGTGGCTTCTTA	1612
OY	1771	GAGTGTCTTACAGATATGAGACAGATGGGTACCAAGTCCATTTGACACCACTGTCTGTGGG	1830
Db	1613	GAGCTCTTACATGATATGATGAGAGAGGGGTATTCAGTCTCTACTGCCACCGTGTGCTATG	1672
OY	1831	GAGGCGTAAGTGCATATGTGTGGAAACAAACACTCTCTCAAGTGTCTTTGTCTCAAGTGT	1890
Db	1673	GAGCGTAACTGTGCTGTGTGCAGTAAACAAAGTGCTGTCAATCTCTTGTGTGGAGTGT	1732
OY	1891	GTGGATCTCTGGTGGGGCCAGAGGCGTGTCAAGGAGCCATTAAGGAAGAGCCCTGGAGAC	1950
Db	1733	CTGGAGAGTGTGTGTGGCGGCACAGCTGACAGATATGCCAAAGCTGCGAGNACCTCTGGAGC	1792
OY	1951	TGCTATCTGTGCGGCAATAAAGGCACTTATGGCTGTGTGCGAAGAGCGGAATATGTGCTT	2010
Db	1793	TGTATATGTGGCTCTCCCTCAAGCGTGTGCATGGGGTCTCTCGACGCAAGAAATATTTGAC	1852
OY	2011	TCTGTACATCAAGATGTTTGGCCAAATACATGACC---AGGATTTTGAAGCCGCCCAAG	2067
Db	1853	ATGCGGCTCTGAGACTTCTTCTACTACTGATCTCTGACCTGGAGAAATTTTGAAGCCACCCAA	1912

OY		2068	GTTTACCACCTGTCGCAGCTGAGAAAGAACGCCATCCCGCTGTCTCTCTTTGAT	2127
Dp		1913	TTGTAACCCAGCAATTCTCGACGCCAAAAGAGGCCCATTTAGAAGTCGTCTCTTTGAT	1972
OY		2128	GGATTTGCTACAGGGCTCTCTGGTGTCTAAGAGCCTGGGCACTCCAAAGTAGAGCCGTACTT	2187
Dp		1973	GGATTTGTCMACGGGGTACTTGCTGTCTCAGAGAGTTGGGTATTTAAAGTGAAGATACATT	2032
OY		2188	GCTTCGAGAGGTGTGTGAGACATCCATCACAGGTGGGATGTGTGGCCGACCAAGGAAATC	2247
Dp		2033	GCCCTCCAAGCTCTGCAGAGTCCATCGTGTGGGAAGCTGTTAAGCATGAAGCCAGATC	2092
OY		2248	ATGTACGTCCGGGAGACCTCCCAGGCTACACAGAAAGATATCCAGAGGTGGGCCCATTC	2307
Dp		2093	AAATATGTCAATGACGTCCGGAAAAATACCAAGAAAAATATTGAAGATGGGGCCCTTC	2152
OY		2308	GACCTGTGATTTGAGGACAGTCCCTCGCAATGACCTCTCCATGTGCAACCTCGCCCGAG	2367
Dp		2153	GACTTGTGTGATTTGTGGAAGCCCATTCGAATGATCTCTTAACGTCAATCTCTCCCGAAA	2212
OY		2368	GGACTTTATGAGGCTACTGGCCGCTCTTTGAGTTCTAACCGCCCTTCGATGATGCG	2427
Dp		2213	GGTTTATATGAGGGACACAGGAAGGCTCTCTCGAGTTTTACCACTTGTAATTAATAC	2272
OY		2428	CGCCCCAAGGAGGAGATGATCCGCCCTTCTCTTGCGCTCTTTGAGATTTGGTGGCATG	2487
Dp		2273	CGCCCCAAGGAGGGCGACAACCGTCAATTTCTTGAGATGTTCCGAATTTGTGGCCATG	2332
OY		2488	GGCGTTTATGTCACAGAGGACATCTCCGATTTTGTAGTCTAACCCCGATGATTTGAC	2547
Dp		2333	AAATGTGATGTCACAGAAAGACATTCAGATTTCTTGCATCTTAACCCAGTATGATGAT	2392
OY		2548	GCCAAAGAAGTGTCTGTGACACACAGGGCCGTTACTTCTGGGGTMACTTCTCTGACATG	2607
Dp		2393	GCCATCAAGGTGTCTGTGTCACACAGGGCCCGGTACTTGTGGGTAACTTAACCGGANTG	2452
OY		2608	AACAGGCTTTGGGATCCACTGTGAATGATTAAGTGTGAGCTGCAGAGTGTCTGTGACAC	2667
Dp		2453	AACAGGCCCCGATGGCTTCAAAGATATAAGTGTGAGTGCAGGAGCTGCTGGGAATTC	2512
OY		2668	GGCAGAATATGACCAATTCAGCAAAAGTAGAGACATTAACACACAGGTCAAACTATPAAAG	2727
Dp		2513	AGTAGGACAGCAAAAGTTAAAGAAAGTGCACGATTAACACACAGTGGAACTCATCAGA	2572
OY		2728	CAGGGCAAAAGCACGACATTTCCCGCTTCATGAAAGAGAAAGAGACATCTGTGTGTC	2787
Dp		2573	CAGGGCAAAAGCACGATTTCCCTGTGTATGATTAAGGAAAGACGAGCTTTGTGTGTC	2632
OY		2788	ACTGAATATGGAAGAGGTGTGTGGCTTCCCGCTCACATACAGACGTCTCCACATAGAC	2847
Dp		2633	ACTGAAGTTCGAAAGGATCTTCGGCTTCCCTGCTCACACAGGACGTGTCCAAATGSGGC	2692
OY		2848	CGCTTGGCGAGACGACGATCTGTGGCGGATCGTGAAGCGTGGCGGTCATCCGACACTC	2907
Dp		2693	CGCGGCGCCCTCGAAGAGCTGTGGGACAGTCTCTGAGGTGTTACCGGTATACACACCTG	2752
OY		2908	TTGCTCCGCTGAAGAAATATTTTGTCTTGTGTGA	2942
Dp		2753	TTTGCCCCCTTGAAGGACTACTTTGGCTGTGAATA	2787
RESULT 15				
AFI51974				
LOCUS	AFI51974	4163 bp	mRNA	linear ROD 08-JUN-2000
DEFINITION	Mus musculus DNA cytosine-specific methyltransferase isoform 6			
ACCESSION	AFI51974			
VERSION	AFI51974.1	GI:8347130		
KEYWORDS	Mus musculus.			
SOURCE	Mus musculus.			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			

[illegible]

1291 CCCATGACGGAAGCCATCTACGAAGTCTCCAGAGTGGCCAGGAGCGGCGGAG 1350
1203 GTTCTTATAGGAAGCCATGTACCAACACTGCGAAGAAAGCCAGGGGTTGAGCTGGCAAG 1262
1351 CTGTTCCAGCTGGCATGACAGTGAAGAGACAGTGGCAAGGCGTGGAGTGCAG 1410
1263 ACCTTCTC-----CAGAGTCTGGAGATCTACTGGAGACAG 1301
1411 AACAGCAATATTTGAATGGGCGCTGGTGGCTTCCAGCCCTCGGGTCTTAAGGCGCTG 1470
1302 CTGAAGCCCATCTGAGTGGGCCCGACGGTGGCTTCAAGCTTCACTGGAGTGAAGGCGCTC 1361
1471 GAGCCACGAAAGAGAGAAATCTTACAGGAAGTTTACAGCCAGATGTGGGTGAG 1530
1362 AAAC-----CAACAAAGCAACCGAGAACMAAAAGTCGAAGACGCAACCAATGAC 1415
1531 CTTGAAGCAGCTGCTTACGCCCCACCCCAAGCCAGCAAGAAACCGAAGACCAACA 1590
1416 TCTGCTGCTTCTGAGTCCCCCAACCAAGCCCTCAAGCAAAATGCTATGGGCGAG 1475
1591 GAGAAACTAAGTCAAGAGATCATTTGATGAGCGCAAGAGGAGCGCTGGTATGAG 1650
1476 GACCGAGGGGAGATGAGAGAGC-----CGAAGAGGATGGCTTGGA 1520
1651 GTGCGCGAAGTGAAGAAACATGAGAGACATTTGTATCTCATGTGGAGGCTCAATGTC 1710
1521 GTACCAACAAACAGGGCAATCTGGAAGACCGCTTTGTCTGTGGAAGAAACCAACCT 1580
1711 ACCCTGAGACAGCCTCTTCAATGAGAGATGCGCAACTGTGAACACTGCTTCTTG 1770
1581 GTGCTCTTCCACCTCTTGAAGGAGGCGCTCTGCAAGATGGCGGAGTCTGCTCTTA 1640
1771 GAGTGTCTTACAGTATGACAGAGATGGTACCAGCTCTTATGACCACTGCTGGG 1830
1641 GAGCTCTTCAATATGATGAGAGAGCGCTATCAGTCTACTGCTGAGTGTGAG 1700
1831 GGGCGTGAAGTCTCATGTGGAGAACACMACACTGTGAGTGGCTTTTGTGAGATGT 1890
1701 GCGCGTGAAGTCTCTCTGTGCACTACCAAGCTGCTGAGTGTCTGCTGAGATGT 1760
1891 GTGATCTCTTGTGGGCGCAGAGCTCTCAGGACCCATTMAAGAAAGACCCCTGAGAC 1950
1761 CTGAGAGTGTGTGGGCGCAGAGCAGCTGAGATGCAACACTCAGAAACCTGTGAGC 1820
1951 TGCTACATGTGGGCGCATAAAGGACCTATGGGCTCTGCAAGAGCGGAGAGTGGCT 2010
1821 TGCTATATGTGCTCTCAGCGCTGCAATGGGCTCTCGACGCAAGAAATTTGAGC 1880
2011 TCTGACCTCCATATGTTTGTGCAATTAACATGACC--AGGAATTTGACCCCGCAAG 2067
1881 ATGCGGCTCAGAACTTTTCTACATGATCTGACCTGGAAGATTTGAGCCACCAAG 1940
2068 GTTACCCACCTGTGGCAGCTGAGAAAGAAAGCAACCCAGCGTCTCTTGTAT 2127
1941 TTGTACCCACCAATTTCTGCAAGCCAAAGAGGCCCATTAAGTCTGTCTGTAT 2000
2128 GGGATTGTACAGGGCTCTGCTGCTGAAGAGCTGGCATCCAAAGTGCAGCTACAT 2187
2001 GGAATGTCAAGGGGTACTTGTGTCTCAAGAGTGGGTATTAAGTGAAGAAAGTACAT 2060
2188 GCGTCCGAGGTGTGAGAGCTCCATCAGCGTGGCATGTTGCGGACCAAGGAAGATC 2247
2061 GCGTCCGAGGTGTGAGAGCTCCATCAGCGTGGCATGTTMAAGCATGAAGGCAATC 2120
2248 ATGTACGTGGGAGAGTCCGAGAGCTGACAGAAACATATCAGAGAGTGGGCGCATTC 2307
2121 AAATATGTCAATGAGTCCGGAATAATCACAAGAAATATTAAGAGTGGGCGCTTC 2180
2308 GACCTGTGATGGAGAGTCCGAGATGACATCTCATATGTCACACCTGGCGGCAAG 2367
2181 GACTTGTGATGGAGAGCCCATGCAATGATCTCTTAAGTCAATCTGCGCGCAAA 2240
2368 GGACTTTATGAGGAGTACGCGCGCTCTTGTGAGTCTACCGCGCTCTCATGATGCG 2427

2241 GTTATATAGGAGGACAGGAGGCTCTTCTTGAAGTTTACACTGTGAAATTAATAC 2300
2428 CGGCTTAAGAGGAGAGATGCGCCCTTCTTCTGAGCTTTGAGATGTGGGCGATG 2487
2301 CGGCTTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2360
2488 GCGCTTATGACAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2547
2361 AAGTGAATGACAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2420
2548 GCGAAAGAGTGTCTGTGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2607
2421 GCGATCAAGGTGTCTGTCTGCTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2480
2608 AACAGGCTTTGGCATGCACTGTGAATGATGAAGCTGAGAGCTGCAAGAGTCTGAGAC 2667
2481 AACAGGCGCGTATGCTTCAAGAAATGATGAAGCTGAGAGCTGCAAGAGTCTGAGAG 2540
2668 GCGAATATAGCCAGTTCAGCAAAAGTGAAGACATTAACACAGGTCAAACTATTAAG 2727
2541 AGTAGACAGCAAGTTAAAGAAAGTGAAGACATTAACACAGGTCGAAGTCTCATACA 2600
2728 CAGGCGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2787
2601 CAGGCGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2660
2788 ACTGAATGAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2847
2661 ACTGAGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2720
2848 GCGTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2907
2721 GCGGCGGCGCGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2780
2908 TTCGCTCGCTGAAGGAATATTTTGTGTGTGA 2942
2781 TTTGCGCGCTTGAAGGAGTACTTTGCTGTGAATA 2815

Search completed: July 17, 2003, 07:36:48
Job time: 7052.05 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 16, 2003, 22:42:04 ; Search time 7013.06 Seconds

(without alignments)
17408.415 Million cell updates/sec

Title: US-09-720-086-2

Perfect score: 4195

Sequence: 1 gaattccggcgccgcgggtt.....aaaaaaaaaaaaaaaaaaaa 4195

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Geneml: *
1: gb_ba: *
2: gb_hcg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vl: *
30: em_hcg_hum: *
31: em_hcg_inv: *
32: em_hcg_other: *
33: em_htg_mus: *
34: em_htg_dln: *
35: em_htg_rtd: *
36: em_htg_mam: *
37: em_htg_vrt: *
38: em_sy: *
39: em_htg_hum: *
40: em_htgo_mus: *
41: em_htgo_other: *

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4195	100.0	4195	10	AF068626 Mus muscu
2	4182.8	99.7	4338	10	AF151969 Mus muscu
3	4065	96.9	4135	10	AF068627 Mus muscu
4	4052.8	96.6	4278	10	AF151970 Mus muscu
5	3942.8	94.0	4223	10	AF151973 Mus muscu
6	3812.8	90.9	4163	10	AF151974 Mus muscu
7	3807	90.8	4006	10	AY078427 Mus muscu
8	3794.8	90.5	4149	10	AF151972 Mus muscu
9	3677	87.7	3946	10	AF068628 Mus muscu
10	3664.8	87.4	4089	10	AF151971 Mus muscu
11	3554.8	84.7	4034	10	AF151975 Mus muscu
12	3424.8	81.6	3974	10	AF151976 Mus muscu
13	1815.4	43.3	4335	9	AF331857 Homo sapi
14	1812.2	43.2	4145	9	AF156488 Homo sapi
15	1697.2	40.5	4267	9	AF176228 Homo sapi
16	1459.8	34.8	244329	2	AC107644 Mus muscu
17	1335.8	31.8	3897	9	AF156487 Homo sapi
18	1072.6	25.6	3017	9	AK001191 Homo sapi
19	858.4	20.5	167568	2	AC111734 Rattus no
20	762	18.2	3005	9	AF067972 Homo sapi
21	762	18.2	4258	9	AF331856 Homo sapi
22	761	18.1	4094	10	BC007466 Mus muscu
23	759.4	18.1	4192	10	AF068625 Mus muscu
24	696	16.6	2191	9	AK025230 Homo sapi
25	656.2	15.6	2127	9	AF129267 Homo sapi
26	548.4	13.1	2077	6	AR129189 Sequence
27	537.8	12.8	176697	2	AC112586 Rattus no
28	525.2	12.5	2848	5	AF135438 Dario fer
29	520.6	12.4	2057	9	AF129268 Homo sapi
30	516.2	12.3	1758	9	BC018214 Homo sapi
31	435.6	10.4	188936	2	AC128509 Rattus no
32	422.6	10.1	2008	9	AF129269 Homo sapi
33	411.4	9.8	168651	9	AC009474 Homo sapi
34	402.4	9.6	123936	2	AC120824 Rattus no
35	266	6.3	119630	2	AC112040 Rattus no
36	248.2	5.9	110000	2	AL354832 Homo sapi
37	248.2	5.9	118899	9	HS1085F17 Human DNA
38	196.4	4.7	49478	2	AC091678 Mus muscu
39	196.4	4.7	244329	2	AC107644 Mus muscu
40	191.8	4.6	49478	2	AC091678 Mus muscu
41	188.6	4.5	225045	2	AL833803 Mus muscu
42	187	4.5	249245	2	AC122356 Mus muscu
43	181	4.3	210269	2	AC116459 Mus muscu
44	178.8	4.3	139581	2	AC131350 Rattus no
45	178.8	4.3	153533	2	AC095199 Rattus no

ALIGNMENTS

RESULT 1
AF068626 4195 bp mRNA linear ROD 06-DEC-1999
LOCUS AF068626
DEFINITION Mus musculus DNA cytosine-5 methyltransferase 3B1 (Dnmt3b) mRNA,
alternatively spliced, complete cds.
ACCESSION AF068626
VERSION AF068626.2 GI:6449469
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 4195)
Okano, M., Xie, S. and Li, E.
Cloning and characterization of a family of novel mammalian DNA

Pred. No. is the number of results predicted by chance to have a

QY 1261 ACTGAGGACAGCTGAGAGCCCATGCTGAGTGGGCCACGCTGGCTTCAAGCCCTACTGG 1320
DB 1261 ACTGAGGACAGCTGAGAGCCCATGCTGAGTGGGCCACGCTGGCTTCAAGCCCTACTGG 1320
QY 1321 GATCGAGGGCTCAAAACCCAAAGAAAGCAACAGTGGTTAATATAGTCGAAAGTGGCTGC 1380
DB 1321 GATCGAGGGCTCAAAACCCAAAGAAAGCAACAGTGGTTAATATAGTCGAAAGTGGCTGC 1380
QY 1381 TTGAGACAGTAGGAACCTTAGAACCCAGAGACGCGAAGAACAAAGTGGAGACCCACAAC 1440
DB 1381 TTGAGACAGTAGGAACCTTAGAACCCAGAGACGCGAAGAACAAAGTGGAGACCCACAAC 1440
QY 1441 CAATGACTCTGCTCTCTAGTCCGCCACCAAGCGCTCTCAAGCAAAATAGCTATGG 1500
DB 1441 CAATGACTCTGCTCTCTAGTCCGCCACCAAGCGCTCTCAAGCAAAATAGCTATGG 1500
QY 1501 CGGGAAGGACCGAGGGGAGATGAGAGACCCGAAACGATGGCTTCTGAAAGTCACAA 1560
DB 1501 CGGGAAGGACCGAGGGGAGATGAGAGACCCGAAACGATGGCTTCTGAAAGTCACAA 1560
QY 1561 CAACAAAGGCAATCTGGAAGACCGCTGTTGCTCTGAGAAAGAAACCTGTGCTCTT 1620
DB 1561 CAACAAAGGCAATCTGGAAGACCGCTGTTGCTCTGAGAAAGAAACCTGTGCTCTT 1620
QY 1621 CCAACCCCTCTTGAAGGGTGGGCTCTGTCAGAGTTGGCGGAGATGGCTTCTAGAGCTCTT 1680
DB 1621 CCAACCCCTCTTGAAGGGTGGGCTCTGTCAGAGTTGGCGGAGATGGCTTCTAGAGCTCTT 1680
QY 1681 CTACATGTATGATGAGAGACGCGCTATGACCTTACTGACCGTGTGCTGAGGGCCGTGA 1740
DB 1681 CTACATGTATGATGAGAGACGCGCTATGACCTTACTGACCGTGTGCTGAGGGCCGTGA 1740
QY 1741 ACAGCTGCTGTGAGTAACAAGCCTGCTGACAGTGGCTGTGTGTGAGAGTCTGAGAGT 1800
DB 1741 ACAGCTGCTGTGAGTAACAAGCCTGCTGACAGTGGCTGTGTGTGAGAGTCTGAGAGT 1800
QY 1801 GCTGTGTGGGCGCAGGCGACACCTGAGATGCCAAGCTCAGAGAACCTGAGAGCTCTATAT 1860
DB 1801 GCTGTGTGGGCGCAGGCGACACCTGAGATGCCAAGCTCAGAGAACCTGAGAGCTCTATAT 1860
QY 1861 GTGCTCTCCACAGGCGCTGAGGCTGCTCCGACGACGAAATGGAACATGAGCCCT 1920
DB 1861 GTGCTCTCCACAGGCGCTGAGGCTGCTCCGACGACGAAATGGAACATGAGCCCT 1920
QY 1921 GCAAGACTTCTTCACTGATCTGATCTGACCTGGAAGATTTGAGCCACCAAGTGTACC 1980
DB 1921 GCAAGACTTCTTCACTGATCTGATCTGACCTGGAAGATTTGAGCCACCAAGTGTACC 1980
QY 1981 AGCAATTCCTGACGCCAAAGAGAGGCCATTAGAGTCTGTCTGTGTTGATGGAATTCG 2040
DB 1981 AGCAATTCCTGACGCCAAAGAGAGGCCATTAGAGTCTGTCTGTGTTGATGGAATTCG 2040
QY 2041 AACGGGCTACTTGGTGTCTCAAGAGTGGGTATTAAAGTGAAGAAATGACATTTGCCGA 2100
DB 2041 AACGGGCTACTTGGTGTCTCAAGAGTGGGTATTAAAGTGAAGAAATGACATTTGCCGA 2100
QY 2101 AGTCTGTGACAGATCTGCTGAGGAGCTGTTAAACATGAAGGCCAGATCAAAATATGT 2160
DB 2101 AGTCTGTGACAGATCTGCTGAGGAGCTGTTAAACATGAAGGCCAGATCAAAATATGT 2160
QY 2161 CAATGACGTCGGAAGAAATCAACAAGAAATATTGAAGAGTGGGGCCGCTTCACTTGT 2220
DB 2161 CAATGACGTCGGAAGAAATCAACAAGAAATATTGAAGAGTGGGGCCGCTTCACTTGT 2220
QY 2221 GATTGGTGGAAAGCCCATGCAATGATCTCTAACGTCATCTGCCCCGCAAGGTTTATA 2280
DB 2221 GATTGGTGGAAAGCCCATGCAATGATCTCTAACGTCATCTGCCCCGCAAGGTTTATA 2280
QY 2281 TGAGGGGACAGGAAGGCTCTTCTGAGTTTACCACTGTGATATATACCCGCCCAA 2340
DB 2281 TGAGGGGACAGGAAGGCTCTTCTGAGTTTACCACTGTGATATATACCCGCCCAA 2340
QY 2341 GGAAGGGGACAGGCTCATTTCTTGATGTTGAGAAATGTTGGCCATGAAAGTAA 2400

DB 2341 GGAAGGGGACAGGCTCATTTCTCTGATGTGTGAGATGTGTGGCCATGAAAGTAA 2400
QY 2401 TGACAAAGAAACATCTCAGATTCCTGGCATGTGAACCCAGATGATGATGCCATCAA 2460
DB 2401 TGACAAAGAAACATCTCAGATTCCTGGCATGTGAACCCAGATGATGATGCCATCAA 2460
QY 2461 GGTGCTGCTGCTACAGAGGCGCGGATCTTGTGGGTAACTTACCCGGAATGAACAGGCC 2520
DB 2461 GGTGCTGCTGCTACAGAGGCGCGGATCTTGTGGGTAACTTACCCGGAATGAACAGGCC 2520
QY 2521 CGTGATGGCTTCAAGATGATGATAGCTGAGCTCAGAGACTGCTGGAGTTCACTAGAGAC 2580
DB 2521 CGTGATGGCTTCAAGATGATGATAGCTGAGCTCAGAGACTGCTGGAGTTCACTAGAGAC 2580
QY 2581 AGCAAAAGTTAAAGAAAGTGCAGACATTAACCCAGTGCAGATCTCATCAGAGGSCAA 2640
DB 2581 AGCAAAAGTTAAAGAAAGTGCAGACATTAACCCAGTGCAGATCTCATCAGAGGSCAA 2640
QY 2641 AAACAGCTTTCCTGTAGTCATGATGGAAGAGAGAGCTTTGTGGTGCATGAGCT 2700
DB 2641 AAACAGCTTTCCTGTAGTCATGATGGAAGAGAGAGCTTTGTGGTGCATGAGCT 2700
QY 2701 CGAAAGATCTTGGGCTTCCCTGCTCACTACACGAGCTGTCCAAACATGGGCCGCGCGC 2760
DB 2701 CGAAAGATCTTGGGCTTCCCTGCTCACTACACGAGCTGTCCAAACATGGGCCGCGCGC 2760
QY 2761 CCGTCAGAGCTGCTGGGCGAGGCTCTGAGTACCGGTCAACACCTGTTGGCCC 2820
DB 2761 CCGTCAGAGCTGCTGGGCGAGGCTCTGAGTACCGGTCAACACCTGTTGGCCC 2820
QY 2821 CTTGAAGAGACTTTCCTGTGTAATGTTCTACCCAGAGATGGGGAGCTCTGGTCA 2880
DB 2821 CTTGAAGAGACTTTCCTGTGTAATGTTCTACCCAGAGATGGGGAGCTCTGGTCA 2880
QY 2881 GCCAGTCCCGAGATCAACCCCTCCCTAAGGACACTACCTGCTTTTAACTTACCTAAC 2940
DB 2881 GCCAGTCCCGAGATCAACCCCTCCCTAAGGACACTACCTGCTTTTAACTTACCTAAC 2940
QY 2941 TGTGTGGGCTCTCACTACATCTGATCTGATCTTCTCTGCTGCTGAGTGGAGACAGCTC 3000
DB 2941 TGTGTGGGCTCTCACTACATCTGATCTGATCTTCTCTGCTGCTGAGTGGAGACAGCTC 3000
QY 3001 CTGGCCCTTGAAGGGAGGCCCGGCTGCTCCCTGCTGAGACACTCAGACCTGAGCTGCT 3060
DB 3001 CTGGCCCTTGAAGGGAGGCCCGGCTGCTCCCTGCTGAGACACTCAGACCTGAGCTGCT 3060
QY 3061 TAGAGTAGCCCGCATGCTGATCTTCTTACCTGAACTTTAAACTTGAAGTAG 3120
DB 3061 TAGAGTAGCCCGCATGCTGATCTTCTTACCTGAACTTTAAACTTGAAGTAG 3120
QY 3121 GTAGTAAGATGCTTTCTTTTACCCTCTGAGTTTATCACCAAGAGTGGCTAAGAT 3180
DB 3121 GTAGTAAGATGCTTTCTTTTACCCTCTGAGTTTATCACCAAGAGTGGCTAAGAT 3180
QY 3181 ACCAAAAAACAACAAAAACAGAAACAAAAAACAACAAAAAACCCTCAACAGCTCTGTA 3240
DB 3181 ACCAAAAAACAACAAAAACAGAAACAAAAAACAACAAAAAACCCTCAACAGCTCTGTA 3240
QY 3241 GTACTCAGCTTCACTGCAAAATACCTTGAATTTGTGTTTAACTAACCCGTGCTCA 3300
DB 3241 GTACTCAGCTTCACTGCAAAATACCTTGAATTTGTGTTTAACTAACCCGTGCTCA 3300
QY 3301 CATTTGCTGGAGATGCTATTGTGAATGTGGGCTCAGATGAGCAAGTCAAGGGCCAAA 3360
DB 3301 CATTTGCTGGAGATGCTATTGTGAATGTGGGCTCAGATGAGCAAGTCAAGGGCCAAA 3360
QY 3361 AAAAATTCCTCCCTTCCCTCCAGAGATTTGAAGATGATGATGATGATGATGATGATGAT 3420
DB 3361 AAAAATTCCTCCCTTCCCTCCAGAGATTTGAAGATGATGATGATGATGATGATGATGAT 3420
QY 3421 TGACACCTTCCCTTCTGTTGTTCAAGGGCTGAGTCTGTTGCTTGTGAGCTTCC 3480

Db	3421	TGCGACGTCGCCCTTCCTTGGTAAAGAGGCTGAACCTCCTGTTGGTCTGTACATTTCC	3480
Oy	3481	CAGGATGATGATGTACAGCAGGGATGACATGCACACCTTTAGGGCGTTTCCCTGGCAGGG	3540
Db	3481	CAGGATGATGATGTACAGCAGGGATGACATGCACACCTTTAGGGCGTTTCCCTGGCAGGG	3540
Oy	3541	CCCATGTGGCTAGTCCCTCCAGAACTGGAGTGTGAATGTTTGAAGCTCAGAAAGGCTGG	3600
Db	3541	CCCATGTGGCTAGTCCCTCCAGAACTGGAGTGTGAATGTTTGAAGCTCAGAAAGGCTGG	3600
Oy	3601	TGGAATGACCTCTTCCACAGTGTGAGGGATTCAGAACAGAAAGCTTTGGGAAATTCATTC	3660
Db	3601	TGGAATGACCTCTTCCACAGTGTGAGGGATTCAGAACAGAAAGCTTTGGGAAATTCATTC	3660
Oy	3661	CCCACTCCCTCTTTGCCAAATGAGGGGCCAGCTCCCAACAGCTCAGGTCCCGAAGCC	3720
Db	3661	CCCACTCCCTCTTTGCCAAATGAGGGGCCAGCTCCCAACAGCTCAGGTCCCGAAGCC	3720
Oy	3721	CTAGTTCCTCATAGAAAGTATGAGACAGAAAGCAATCGTCCCTTCTGACACTGTT	3780
Db	3721	CTAGTTCCTCATAGAAAGTATGAGACAGAAAGCAATCGTCCCTTCTGACACTGTT	3780
Oy	3781	TGGGGAAGTACATGAAACCTCTCGAGATGTAAACCTTTTACCACAGTATGAT	3840
Db	3781	TGGGGAAGTACATGAAACCTCTCGAGATGTAAACCTTTTACCACAGTATGAT	3840
Oy	3841	GTGTTTAAAGGGCTGCTTTTAAAGGGCATACAGAGATGAAGAAAGTGCATTTAC	3900
Db	3841	GTGTTTAAAGGGCTGCTTTTAAAGGGCATACAGAGATGAAGAAAGTGCATTTAC	3900
Oy	3901	AAAGGCCATGTAAAGCTTTTAAACACCTTTTACTATTAACAGTGCTATTTATAGA	3960
Db	3901	AAATGCCATGTAAAGCTTTTAAACACCTTTTACTATTAACAGTGCTATTTATAGA	3960
Oy	3961	AGCAGACAAACCTTTTAAAGCTCAGACTCTATTTATGTTTACCATTATTTT	4020
Db	3961	AGCAGACAAACCTTTTAAAGCTCAGACTCTATTTATGTTTACCATTATTTT	4020
Oy	4021	GTAACTGGCAAGGTGTGGGCTTTGTAACTTCACAGTGTGGGAGAGACTGCTCTTT	4080
Db	4021	GTAACTGGCAAGGTGTGGGCTTTGTAACTTCACAGTGTGGGAGAGACTGCTCTTT	4080
Oy	4081	CACACGTTTGTCTCCACTGTGTTCTAATTTTAAAGTGCAGAAATGACAGATCCCACT	4140
Db	4081	CACACGTTTGTCTCCACTGTGTTCTAATTTTAAAGTGCAGAAATGACAGATCCCACT	4140
Oy	4141	TTACCTTTCTGCTTGAATTAAAGTGTATTTCTATAAAAAAAAAAAAAAAAAAAAAA	4195
Db	4141	TTACCTTTCTGCTTGAATTAAAGTGTATTTCTATAAAAAAAAAAAAAAAAAAAAAA	4195

RESULT 2	AF151969	4338 bp	musculus	RNA	linear	ROD 08-JUN-2000
NAME	AF151969					
DEFINITION	Mus musculus DNA cytosine-specific methyltransferase isoform 1 (Dnmt3b) mRNA, complete cds.					
ACCESSION	AF151969					
VERSION	AF151969.1					
KEYWORDS	GI:8347117					
SOURCE	Mus musculus.					
ORGANISM	Mus musculus					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.					
ATTORNS	Yin,B., Chen,Y.T., Zhu,M., Luo,Y.J., Zhu,N., Xu,S.C., Wu,G.Y. and Shen,Y.					
TITLE	Cloning of full-length Dnmt3b cDNA and its alternative splicing isoforms in mouse embryonic tissue					
JOURNAL	Unpublished					
REFERENCE	2 (bases 1 to 4338)					
ATTORNS	Yin,B., Chen,Y.T., Zhu,M., Luo,Y.J., Zhu,N., Xu,S.C., Wu,G.Y. and Shen,Y.					
TITLE	Direct Submission					

JOURNAL

Submitted (17-MAY-1999)

Department of Biochemistry and Molecular Biology, Institution of Basic Medical Sciences, Chinese Academy of Medical Sciences/ Dong Dan san Tiao 5, Beijing City 100005, P.R.China

FEATURES

source

1..4338

/organism="Mus musculus"

/strain="KM"

/db_xref="taxon:10090"

/dev_stage="8-9 day old embryo"

gene

1..4338

/gene="Dmrt3b"

412..2991

/gene="Dmrt3b"

/function="methylates cytosine in DNA"

/note="alternatively spliced"

/codon_start=1

/product="DNA cytosine-specific methyltransferase isoform 1"

CDS

/protein_id="AA074515.1"

/db_xref="gi:8347118"

/translat="MKGDSRLNEEGAGSYECIIVNGNFSDSDTDAKPSVVEV
ALTEVEVCTPTETGRSSRLNREVSLSLANTQMDGDRDVEDDNGSDILMPLE
LRETEEDTRRESFAVYTRNSGTSLEKROASPRITRGROGHNVQY PVEPAT
SRRRRASASPTSPSPASVDPMEEVTPKSVSTPSVDLSODGQEMDITVDASIS I
SETSAOKDDEFGIDGLDWGIKGFSMPMAVYSKATKAKAPGMRWONFGGKGL
OLMPLEMAHGKFKTJGIEGLKPNKOPVNVKSVKRSDSRLNLEPRRNSKRRPTT
DNASSEPPRLKTNVSGGDRGDEDEERENMAEVEVNNKNGLEDRCLSGKKNPV
FHLVEAGACOSCRRDRLLEIPLYMDGOSVCTCGEGRLLGNSGSCGPFVECE
LELVAGAGAEADAKLOEPMWSCYMLPDRCHGYLRRKROMMNRLDPEFTDPLEE
PKLYPAAPAKRRPRIVLSLPEGLATGYSLDELIVEXYIASVECAISAVGVKK
EGQIKVYNDNRKRTTKNIEMSGPFDLYIGSGSNDLNVNPAKGLYEGTGLPEEF
HLNTRPREKGNRPPEMFENYVAMKVNKDDISRLPACNPMIDALIVSAHARV
FMNLEKGNMRPYMASKNDKLIEDDCLFESRTKLYKQITTTSSITROGKNLPPVY
MNRKDYLMCTLERIGTFPAHNTDVSNNGRARKLLGSNSVPIRRLFAPLKDVTY
ACE"

BASE COUNT

1073 a 1124 c 1173 g 968 t

ORIGIN

Query Match

99.7%; Score 4182.8; DB 10; Length 4338;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 4184; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY

10

GGCGCGGGGTTAAAGCGGCCCAAGTAAACGTAGCGCAGCGCATGGCGCGGAGATTCGCA

69

Db

153

GGCGCGGGGTTAAAGCGGCCCAAGTAAACGTAGCGCAGCGCATGGCGCGGAGATTCGCA

212

OY

70

ACCGGACACTCGCGCGCGCGCGCGCGCGCGCAGAACCCGCGCGCATCGCGCGCGCGCTA

129

Db

213

ACCGGACACTCGCGCGCGCGCGCGCGCGCGCAGAACCCGCGCGCGATCGCGCGCGCGCTA

272

OY

130

CAGCGACCTCCACGACGAGCGCCGCGCTGAGCGCTTGTGCAACACCTCTGGAACCTCAGGTATA

189

Db

273

CAGCGACCTCCACGACGAGCGCCGCGCTGAGCGCTTGTGCAACACCTCTGGAACCTCAGGTATA

332

OY

190

TACCTTTCACGAGCGCGGATCTCCCTCCCTCCCAATCCATAGTCCCTGGAGCAAAATCCAG

249

Db

333

TACCTTTCACGAGCGCGGATCTCCCTCCCTCCCAATCCATAGTCCCTGGAGCAAAATCCAG

392

OY

250

GACCTCTTTCACGAAACAAATGAAGGGGAGACACACATCTCGATGATAAAGAGGTGC

309

Db

393

GACCTCTTTCACGAAACAAATGAAGGGGAGACACACATCTCGATGATAAAGAGGTGC

452

OY

310

CAGCGGATAGAGAGATGATATTCGTATTATGGAACTCAAGTACACAGCTCCAGACAC

369

Db

453

CAGCGGATAGAGAGATGATATTCGTATTATGGAACTCAAGTACACAGCTCCAGACAC

512

OY

370

GAGGATGCTCCTCACCCCGACGCTTGGAGGCAATCTGCACAGCCATCTGCACACC

429

Db

513

GAGGATGCTCCTCACCCCGACGCTTGGAGGCAATCTGCACAGCCATCTGCACACC

572

OY

430

AGACACGAGGCGCGGAGTCAAGCTCCCGGCTGTCTTAAGAGGAGGTCTTCACGCTTCT

489

|||||
Db 573 AGAGACGAGAGCGCGAGCTCAACCTCCGGCTGTCTAAGAGGAGGCTCTCCAGCTTCT 632
Qy 490 GAATTAACGAGACACTGACAGAGATGAGACAGATGATGAATGATGATGGA 549
Db 633 GAATTAACGAGACACTGACAGAGATGAGACAGATGATGAATGATGATGGA 692
Qy 550 TGCGTCTGATATTCTAATGCAAAAGCTCACCGGTGACACCAAGGACACAGAGGCGCTC 609
Db 693 TGCGTCTGATATTCTAATGCAAAAGCTCACCGGTGACACCAAGGACACAGAGGCGCTC 752
Qy 610 TGAAGCCCGGCTGTCCGAAACCCGACATAGCAATGAGACCTCCAGCTTGGAGGCAAG 669
Db 753 TGAAGCCCGGCTGTCCGAAACCCGACATAGCAATGAGACCTCCAGCTTGGAGGCAAG 812
Qy 670 AGCTCCCGCAGATCAACCCGAGGTGCGGAGGCGCCACCATGTGACAGAGTACCTGT 729
Db 813 AGCTCCCGCAGATCAACCCGAGGTGCGGAGGCGCCACCATGTGACAGAGTACCTGT 872
Qy 730 GGAATTTCCGGCTACAGGTCTCGAGAGCTGAGCATGCTTCCAGCAACACGCGCATG 789
Db 873 GGAATTTCCGGCTACAGGTCTCGAGAGCTGAGCATGCTTCCAGCAACACGCGCATG 932
Qy 790 GTCAATCCCTGCGAGCGCTGCACTTCAATGAAAGTGAACACTTAAGCGCTGATACCC 849
Db 933 GTCAATCCCTGCGAGCGCTGCACTTCAATGAAAGTGAACACTTAAGCGCTGATACCC 992
Qy 850 ATCAAGTGAATGAGCCAGAGATGAGATGAGAGAGGATGATGCCACACAGGTGATGC 909
Db 993 ATCAAGTGAATGAGCCAGAGATGAGATGAGAGAGGATGATGCCACACAGGTGATGC 1052
Qy 910 AGAGAGCAGAGATGAGAGACAGACAGATATCAGATGATTAAGATTTGGATTAAGTGA 969
Db 1053 AGAGAGCAGATATGAGAGACAGACAGATATCAGATGATTAAGATTTGGATTAAGTGA 1112
Qy 970 CTTCTGTGGGAAAAGATCAAGGGCTTCTCTGTGGCTGCCATGCTGTCTCTGAA 1029
Db 1113 CTTCTGTGGGAAAAGATCAAGGGCTTCTCTGTGGCTGCCATGCTGTCTCTGAA 1172
Qy 1030 AGCCACCTCCAAGGAGAGGAGCCATGCGCGAAATCCGCTGGTAAAGTGTGATG 1089
Db 1173 AGCCACCTCCAAGGAGAGGAGCCATGCGCGAAATCCGCTGGTAAAGTGTGATG 1232
Qy 1090 CAAGTTTCTGAGATCTCTGCTGACAAATGCTGCTCTGGGCTGTTCAAGCAGCACTT 1149
Db 1233 CAAGTTTCTGAGATCTCTGCTGACAAATGCTGCTCTGGGCTGTTCAAGCAGCACTT 1292
Qy 1150 TAATCTGGCTACTTCAATAAGCTGTTCTTATAGAGGCGCATGACACTCTGGA 1209
Db 1293 TAATCTGGCTACTTCAATAAGCTGTTCTTATAGAGGCGCATGACACTCTGGA 1352
Qy 1210 GAAAGCCAGGGTTGAGCTGGCAAGACCTTCTCCAGAGTCTGAGAGTACAGTGA 1269
Db 1353 GAAAGCCAGGGTTGAGCTGGCAAGACCTTCTCCAGAGTCTGAGAGTACAGTGA 1412
Qy 1270 CCAAGTGAAGCCATGCTGAGTGGGCCAAGTGGCTTCAAGCCTACTGGGATCGAGG 1329
Db 1413 CCAAGTGAAGCCATGCTGAGTGGGCCAAGTGGCTTCAAGCCTACTGGGATCGAGG 1472
Qy 1330 CCTCAAAACCAAGAGACCAAGTGTATTAAATAGTCGAGAGTGGCTGTTCAGACAG 1389
Db 1473 CCTCAAAACCAAGAGACCAAGTGTATTAAATAGTCGAGAGTGGCTGTTCAGACAG 1532
Qy 1390 TAGGACCTTAGAACCCGAGAGCGGAGAACAAAGTCGAAGAGCAGCAACCAATGATC 1449
Db 1533 TAGGACCTTAGAACCCGAGAGCGGAGAACAAAGTCGAAGAGCAGCAACCAATGATC 1592
Qy 1450 TGCTGCTTCTGAGTCCCGCCAGCCAGCGCTCAAGCAATAGCTATGGCGGGAAGA 1509
Db 1593 TGCTGCTTCTGAGTCCCGCCAGCCAGCGCTCAAGCAATAGCTATGGCGGGAAGA 1652
Qy 1510 CCGAGGGAAGATGAGAGAGCGAGAAAGGATGCTTCTGAGTCAACCAACCAAGG 1569
|||||

Db 1653 CCGAGGGAAGATGAGAGAGCGAGAACGATGCTTCTGAGTCAACCAACCAAGG 1712
Qy 1570 CAATCTGGAAGACCGCTGTTGTCTGTGGAAGAAAGACCCCTGTCTCCACCCCT 1629
Db 1713 CAATCTGGAAGACCGCTGTTGTCTGTGGAAGAAAGACCCCTGTCTCCACCCCT 1772
Qy 1630 CTTTGAAGGTTGGCTGTGCTGAGAGTCCCGGAGTGGCTTCTAGAGCTCTTCAATGTA 1689
Db 1773 CTTTGAAGGTTGGCTGTGCTGAGAGTGGCGGAGTGGCTTCTAGAGCTCTTCAATGTA 1832
Qy 1690 TGATGAGAGCGGCTATGATCTTACCTGACCGTGTCTGTGAGGCGCGTGAAGTGTCT 1749
Db 1833 TGATGAGAGCGGCTATGATCTTACCTGACCGTGTCTGTGAGGCGCGTGAAGTGTCT 1892
Qy 1750 GTGAGTAACCAAGCTGTGCTGCAATGCTTCTGTGTGAGTGTGAGAGTGTGAGG 1809
Db 1893 GTGAGTAACCAAGCTGTGCTGCAATGCTTCTGTGTGAGTGTGAGAGTGTGAGG 1952
Qy 1810 CGCAGGCACAGCTGAGATGCCAAGCTGACAGGAACCCCTGGAGTGTATATGTCTCC 1869
Db 1953 CGCAGGCACAGCTGAGATGCCAAGCTGACAGGAACCCCTGGAGTGTATATGTCTCC 2012
Qy 1870 TCAAGCTGCCATGGGCTCTCCGACGAGGAAGATTTGGAACATGCGCTGCAAGACTT 1929
Db 2013 TCAAGCTGCCATGGGCTCTCCGACGAGGAAGATTTGGAACATGCGCGCTGCAAGACTT 2072
Qy 1930 CTTCACTACATCTCTGACCTGGAAGATTTGACCACCCAGTTGTAACCAAGATTC 1989
Db 2073 CTTCACTACATCTCTGACCTGGAAGATTTGACCACCCAGTTGTAACCAAGATTC 2132
Qy 1990 TGCAGCCAAAAGAGAGCCCATTAAGTCTCTCTCTTTGATGGAATTCGACAGGGTA 2049
Db 2133 TGCAGCCAAAAGAGAGCCCATTAAGTCTCTCTCTTTGATGGAATTCGACAGGGTA 2192
Qy 2050 CTTGTGCTCAAGAGTGGGTATTAAGTGAAGATGACTTCCCTCGAAGTCTGTC 2109
Db 2193 CTTGTGCTCAAGAGTGGGTATTAAGTGAAGATGACTTCCCTCGAAGTCTGTC 2252
Qy 2110 AGATGCAATGCTGTGGGAACCTGTAAGCATGAAGGCGCATGAATATGTCATGACGT 2169
Db 2253 AGATGCAATGCTGTGGGAACCTGTAAGCATGAAGGCGCATGAATATGTCATGACGT 2312
Qy 2170 CCGGAAAATCAACGAAGAAAATTTGAAGAGTGGGCGCTTCGACTTGTGATTTGGTGG 2229
Db 2313 CCGGAAAATCAACGAAGAAAATTTGAAGAGTGGGCGCTTCGACTTGTGATTTGGTGG 2372
Qy 2230 AAGCCATGCAATATCTCTTAACGTCATCTGCGCCGCAAGAGTTATATGAGGCGAC 2289
Db 2373 AAGCCATGCAATATCTCTTAACGTCATCTGCGCCGCAAGAGTTATATGAGGCGAC 2432
Qy 2290 AGGAAGGCTCTTCTGAGTTTACCACTTCTGAATTAATCCGCGCCCAAGGAGGCGGA 2349
Db 2433 AGGAAGGCTCTTCTGAGTTTACCACTTCTGAATTAATCCGCGCCCAAGGAGGCGGA 2492
Qy 2350 CAACCGTCAATTTCTGAGATGTCGAGAAATGTTGTGGCCATGAAAGTGAACAGAA 2409
Db 2493 CAACCGTCAATTTCTGAGATGTCGAGAAATGTTGTGGCCATGAAAGTGAACAGAA 2552
Qy 2410 AGACATCTCAAGATTTCTGGCATGTACCCAGTATGATCATGCTCAAGGTGTCTGC 2469
Db 2553 AGACATCTCAAGATTTCTGGCATGTACCCAGTATGATCATGCTCAAGGTGTCTGC 2612
Qy 2470 TGCTCACAGGCGCGGATCTCTGAGGATTAACCTCCGAGATGAACAGGCGCGATGGC 2529
Db 2613 TGCTCACAGGCGCGGATCTCTGAGGATTAACCTCCGAGATGAACAGGCGCGATGGC 2672
Qy 2530 TTCAAGAGATGATAGCTGAGCTGACAGACTGCTGAGATTCATGAGACAGCAAGTT 2589
Db 2673 TTCAAGAGATGATAGCTGAGCTGACAGACTGCTGAGATTCATGAGACAGCAAGTT 2732
Qy 2590 AAGAAAGTGCAGCAATTAACCAAGTGCAGTCCATCAAGAGGCAAAACAGCT 2649
Db 2733 AAGAAAGTGCAGCAATTAACCAAGTGCAGTCCATCAAGAGGCAAAACAGCT 2792
|||||

QY	265	TTTCCCTGTAGTATGAAATGCGAAGAGACGTTTGTGTGACATGAGCTGCAAAAGAT	2708
Db	2793	TTTCCCTGTAGTATGAAATGCGAAGAGACGTTTGTGTGACATGAGCTGCAAAAGAT	2855
QY	2710	CTTGGGCTCCCGTCATACATACAGAGACGTGTCCAAATGAGGCGGGCCGCGTGTAGA	2768
Db	2853	CTTGGGCTCCCGTCATACATACAGAGAGGTGTCCAAATGAGGCGGGCCGCGTGTAGA	2912
QY	2770	GCTGCTGGGAGGTCTGTGAGTGTACCGGTATACAGACCTGTPTTGCCCTTTGAGGA	2829
Db	2913	GCTGCTGGGAGGTCTGTGAGTGTACCGGTATACAGACCTGTPTTGCCCTTTGAGGA	2972
QY	2830	CTACTTTGCTGTGAATAGTTCTTACCCAGAGACTGTGGTGTACAGAGCACTGTGC	2889
Db	2973	CTACTTTGCTGTGAATAGTTCTTACCCAGAGACTGTGGTGTACAGAGCACTGTGC	3032
QY	2890	CAGAGTACACCCCTCCCTGAAGGACCTCACCTCTCCCTTTTATAGCTACACTGTGTGGG	2949
Db	3033	CAGAGTACACCCCTCCCTGAAGGACCTCACCTCTCCCTTTTATAGCTACACTGTGTGGG	3092
QY	2950	CCTACATCACTACTTACCTACAGTTTCTCTGCTGCATGTGGAGACAGACCTCTGTGGCCCT	3009
Db	3093	CCTACATCACTACTTACCTACAGTTTCTCTGCTGCATGTGGAGACAGACCTCTGTGGCCCT	3152
QY	3010	GCAGGGAGGAGCCGGGTGCTCCCTCGGTGTACAGCTGCAGACCTGTGGTGTAGATAGC	3069
Db	3153	GCAGGGAGGAGCCGGGTGCTCCCTCGGTGTACAGCTGCAGACCTGTGGTGTAGATAGC	3212
QY	3070	CCGGCATGTGTGCATAGTTTCTTACCGCTGAACCTTAAACTTGAAGTGTAGTATAGA	3129
Db	3213	CCGGCATGTGTGCATAGTTTCTTACCGCTGAACCTTAAACTTGAAGTGTAGTATAGA	3272
QY	3130	TGGCTTTCTTTTACCTCGTGTATTTTACATACAGAGTATGGCTATAGTATACAAAAA	3189
Db	3273	TGGCTTTCTTTTACCTCGTGTATTTTACATACAGAGTATGGCTATAGTATACAAAAA	3332
QY	3190	ACAAACAAAACAGAAACAAAAAACAACAAAAAAGCTCAACAGCTCTTATGACTACAG	3249
Db	3333	ACAAACAAAACAGAAACAAAAAACAACAAAAAAGCTCAACAGCTCTTATGACTACAG	3392
QY	3250	TTTCATGCTGAAATACACTGTGACATTTTGTTTTAAAGTAAACCCGTGCTCCACATTTGCTG	3309
Db	3393	TTTCATGCTGAAATACACTGTGACATTTTGTTTTAAAGTAAACCCGTGCTCCACATTTGCTG	3452
QY	3310	GAGATGCTATTTGTGAATGTGGCTGCAGATGAGCAAGTCAAGGGGCCAAAAAAATTC	3369
Db	3453	GAGATGCTATTTGTGAATGTGGCTGCAGATGAGCAAGTCAAGGGGCCAAAAAAATTC	3512
QY	3370	CCCTCTCCCCCAGAGATATTGGAATGATGTTATGTTAAGTTAAGTCTCTGGACCTT	3429
Db	3513	CCCTCTCCCCCAGAGATATTGGAATGATGTTATGTTAAGTTAAGTCTCTGGACCTT	3572
QY	3430	CCCTTCTCTTTGGTACAGGGGCTGAAATCCTGTGTGCTTGTATACATTTCCAGAGATAT	3489
Db	3573	CCCTTCTCTTTGGTACAGGGGCTGAAATCCTGTGTGCTTGTATACATTTCCAGAGATAT	3632
QY	3490	GATGTACAGCAGGAGATGACATCACACCTTAAAGGCTTTTCCCTGGCAGGGGCCCAATGGG	3549
Db	3633	GATGTACAGCAGGAGATGACATCACACCTTAAAGGCTTTTCCCTGGCAGGGGCCCAATGGG	3692
QY	3550	CTTACTCTCCACGAACTGTGAGTGTGAATGTTTGTGAGCTCAGAAAGGTTGGGTGTGATGGC	3609
Db	3693	CTTACTCTCCACGAACTGTGAGTGTGAATGTTTGTGAGCTCAGAAAGGTTGGGTGTGATGGC	3752
QY	3610	CCCTTCCAGAGTGTAGGATAGCAAGAGGAAAGCTTATGGGAAATTCATTTCCCACTCCC	3669
Db	3753	CCCTTCCAGAGTGTAGGATAGCAAGAGGAAAGCTTATGGGAAATTCATTTCCCACTCCC	3812
QY	3670	TTCTTCCCAATGAGGGGCCAGTCCCAACAGCTCAGTCTCCCAAGAACCCCTCAAGTTCTT	3729
Db	3813	TTCTTCCCAATGAGGGGCCAGTCCCAACAGCTCAGTCTCCCAAGAACCCCTCAAGTTCTT	3872

OY	3730	CATGGAAGCTAGSACCAACAAACATCGTCCCTTATCAGACAGTGTGGGAACT	3789
Db	3873	CATGGAAGCTAGSACCAACAAACATCGTCCCTTATCAGACAGTGTGGGAACT	3932
OY	3790	ACATGAAAACTCTTGGAGATGTTAAAACCTTTTACCCACGATGATGTGTTTTA	3849
Db	3933	ACAGTAAAACTCTTGGAGATGTTAAAACCTTTTACCCACGATGATGTGTTTTA	3992
OY	3850	AGGGGTGCTTTTTTATGGGGCATCAGTGGAGATTAAGAAAAGCTGATTCGAATGCCAT	3909
Db	3993	AGGGGTGCTTTTTTATGGGGCATCAGTGGAGATTAAGAAAAGCTGATTCGAATGCCAT	4052
OY	3910	CGTATGTTTTTAAACACCTTTACCTTAATTCAGTGGCTATTTTTTAAACACACAA	3969
Db	4053	CGTATGTTTTTAAACACCTTTTACCTTAATTCAGTGGCTATTTTTTAAACACACAA	4112
OY	3970	CACTTCTTTTATGACTCTCAGACTCTATTTTCATGTTACATTTTTTTTGTAACTGC	4029
Db	4113	CACTTCTTTTATGACTCTCAGACTCTATTTTTCATGTTACATTTTTTTTGTAACTGC	4172
OY	4030	AAGGTGGGGCTTTTGTAACTTCACAGGTGGGGAGAGACTGCTTGTTCACACGTT	4089
Db	4173	AAGGTGGGGCTTTTGTAACTTCACAGGTGGGGAGAGACTGCTTGTTCACACGTT	4232
OY	4090	GTCGCACCTGTTTCTAATTTTTTATGGGCAAAATATACAGATGGCCAGATTTACCTTC	4149
Db	4233	GTCGCACCTGTTTCTAATTTTTTATGGGCAAAATATACAGATGGCCAGATTTACCTTC	4292
OY	4150	TGCTGTGATTTAAAGTGTATTTCTCTAAAAAATTTTTTTTTTTTTTTTTTTTTT	4195
Db	4293	TGCTGTGATTTAAAGTGTATTTCTCTAAAAAATTTTTTTTTTTTTTTTTTTTTT	4338

RESULT 3	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL	REMARK	FEATURES	
AF068627	4135 bp. mRNA	Mus musculus DNA cytosine-5 methyltransferase 3B2 (Dnmt3b) mRNA, alternatively spliced, complete cds.	AF068627	AF068627			Mus musculus.	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	Okano, M., Xie, S., and Li, E.	Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases	Nat. Genet. 19 (3), 219-220 (1998)	Okano, M., Xie, S., and Li, E.	Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases	Nat. Genet. 19 (3), 219-220 (1998)	Okano, M., Xie, S., and Li, E.	Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases	Nat. Genet. 19 (3), 219-220 (1998)	Okano, M., Xie, S., and Li, E.	Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases	Nat. Genet. 19 (3), 219-220 (1998)	On Nov 18, 1999 this sequence version replaced gi:3327981.	1..4135 /organism="Mus musculus" /db_xref="taxon:10090" /chromosome="7" /map="84.0 cM" 1..4135 /gene="Dnmt3b" 1..4135

/gene="Dmrt3b"
/note="similar to EST sequences deposited in GenBank
Accession Numbers A116594, A119793, A210568,
AA407106, and AA575617"
269. . 2788
CDS

BASE COUNT	1043 a	1048 c	1106 g	938 t
ORIGIN				

Query Match	96.9%	Score 4065;	DB 10;	Length 4135;
Best Local Similarity	98.6%;	Pred. NO. 0;		
Matches 4135;	Conservative	0;	Mismatches	0;
			Indels	60;
			Gaps	1;

OY	I	GAATTCGGGGGGCCGGGGTTAAGGGGCCCAAGTAAAGCTAACGCAGCATTCGGCGCCGA	60
Db	1	GAATTCGGGGGGCCGGGGTTAAGGGGCCCAAGTAAAGCTAACGCAGCATTCGGCGCCGA	60
OY	61	GATTCGGGAACCCGACACTCCGGCCGGCCCGCCGAGACCCGGCGGGATTCGGCGC	120
Db	61	GATTCGGGAACCCGACACTCCGGCCGGCCCGCCGAGACCCGGCGGGATTCGGCGC	120
OY	121	GCCGCGCTACAGCAGCCTCAAGACAGGCCCGCTGAGGCTTGTCGACACCTTGGAAAC	180
Db	121	GCCGCGCTACAGCAGCCTCAAGACAGGCCCGCTGAGGCTTGTCGACACCTTGGAAAC	180
OY	181	TCAGGTATATACCTTTCCAGACGCGGGATCTCCCTCCCATCCATATGTCCTTGGAC	240
Db	181	TCAGGTATATACCTTTCCAGACGCGGGATCTCCCTCCCATCCATATGTCCTTGGAC	240
OY	241	CAATTCACAGGCCCTTTCTTCAGGAAACAATTAAGGAGACACAGCAGACTCTGATGAAGA	300
Db	241	CAATTCACAGGCCCTTTCTTCAGGAAACAATTAAGGAGACACAGCAGACTCTGATGAAGA	300
OY	301	AGAGGGTGCACGCGGGTATGAGAGATGCTATTATGTTAATGSGAATTCAGTGAACAGTC	360
Db	301	AGAGGGTGCACGCGGGTATGAGAGATGCTATTATGTTAATGSGAATTCAGTGAACAGTC	360
OY	361	CTCAGACACGAAGATGCTCCCTCACCCCACTCTTGAGAGCAATCTGCACAGAGCCAGT	420
Db	361	CTCAGACACGAAGATGCTCCCTCACCCCACTCTTGAGAGCAATCTGCACAGAGCCAGT	420
OY	421	CTGCACACCAAGACACGAGAGCCGAGTCAAAGTCCCGGCTGTCTAAGAGGAGGCTTC	480
Db	421	CTGCACACCAAGACACGAGAGCCGAGTCAAAGTCCCGGCTGTCTAAGAGGAGGCTTC	480
OY	481	CAGCCTTGTGAATTACACGACGAGACATACAGGAGATGAGACACAGATGTGATGAAGTGA	540
Db	481	CAGCCTTGTGAATTACACGACGAGACATGACAGGAGATGTGAGACACAGATGTGATGAAGTGA	540
OY	541	TGATGGGAATGCTGTGATATTCTTAATGTCCAAAGCTCACCGGTGAGACCAAGGACACCG	600

D	b		541	TTGATGGAAATGGCTCTGATATTCTTAATGCCAAAGCTCACCCGTGAGACCAAGAACACCAG	600
Q	y		601	GACGGGCTCTTAAAGCCCGGCTGTCCGAAACCCGACATGCAATGGAGACTTCAGCTTGG	660
D	b		601	GACGGGCTCTTAAAGCCCGGCTGTCCGAAACCCGACATGCAATGGAGACTTCAGCTTGG	660
Q	y		661	GAGCGAAAGACCTTCCCGCAGAAATCACCAGAGTGGCGCAGGGCCGCCACCATGTCACGA	720
D	b		661	GAGCGAAAGACCTTCCCGCAGAAATCACCAGAGTGGCGCAGGGCCGCCACCATGTCACGA	720
Q	y		721	GTACCCCTGGAGTTTCCGGCTTACCAGGTCTCCGAGAGCTGAGCATGTCCTTCAGCAAG	780
D	b		721	GTACCCCTGGAGTTTCCGGCTTACCAGGTCTCCGAGAGCTGAGCATGTCCTTCAGCAAG	780
Q	y		781	CACGCGCATGTCATCCTCCCTGCGAGCGTGCATTTATGGAAAGAGACACCTTAAGAGCGT	840
D	b		781	CACGCGCATGTCATCCTCCCTGCGAGCGTGCATTTATGGAAAGAGACACCTTAAGAGCGT	840
Q	y		841	CAGTACCCCATCAGTTGACTTGAGCCAGATGGAGATCAGAGGGGTATGATTACACACA	900
D	b		841	CAGTACCCCATCAGTTGACTTGAGCCAGATGGAGATCAGAGGGGTATGATTACACACA	900
Q	y		901	GGTGGATCCAGAGACAGATGGAGACACAGACAGATATAGATATGATTAAGAGTTTGG	960
D	b		901	GGTGGATCCAGAGACAGATGGAGACACAGACAGATATAGATATGATTAAGAGTTTGG	960
Q	y		961	AATAGTACCTCCGTGGGGAAACATCAAGGGCTTCCTCGTGGCCCTGCGCATGTTGGT	1020
D	b		961	AATAGTACCTCCGTGGGGAAACATCAAGGGCTTCCTCGTGGCCCTGCGCATGTTGGT	1020
Q	y		1021	GTCTCGAAAGCCACCTCCAAAGCAGACAGGCCATGCCCGGAAATGCGCTGGGTACAGTGGTT	1080
D	b		1021	GTCTCGAAAGCCACCTCCAAAGCAGACAGGCCATGCCCGGAAATGCGCTGGGTACAGTGGTT	1080
Q	y		1081	TGTGTATGGCAAGTTTCTTGAGATCTCTGCTGACAAATGTGGTGGCTGTGGGCTGTGTAG	1140
D	b		1081	TGTGTATGGCAAGTTTCTTGAGATCTCTGCTGACAAATGTGGTGGCTGTGGGCTGTGTAG	1140
Q	y		1141	CCAGACATTTAATCTGGCTACCTTCATTAAGCTGTCTTATAGGAAGGCATATACCA	1200
D	b		1141	CCAGACATTTAATCTGGCTACCTTCATTAAGCTGTCTTATAGGAAGGCATATATACCA	1200
Q	y		1201	CACCTCGAGAAAGCCAGAGGTTGAGCTGGCAAGACCTTCTCCACAGTCTCTGAGAGTCTC	1260
D	b		1201	CACCTCGAGAAAGCCAGAGGTTGAGCTGGCAAGACCTTCTCCACAGTCTCTGAGAGTCTC	1260
Q	y		1261	ACTGGAGACACAGCTGAAGCCCATGCTGGAGTGGGCGCCACGGTGGCTTCAAGCCTACATGG	1320
D	b		1261	ACTGGAGACACAGCTGAAGCCCATGCTGGAGTGGGCGCCACGGTGGCTTCAAGCCTACATGG	1320
Q	y		1321	GATCGAGGGCCCTCAAAACCCAAACAGAACCAACCATGTTAATAAGTGGAAAGTCCGTGC	1380
D	b		1321	GATCGAGGGCCCTCAAAACCCAAACAGAACCAACCATGTTAATAAGTGGAAAGTCCGTGC	1380
Q	y		1381	TTTCAGACAGTAGGAATTAGAAACCCAGAGACGCGCAGACAAAATTCGAAGCAGCAAC	1440
D	b		1381	TTTCAGACAGTAGGAATTAGAAACCCAGAGACGCGCAGACAAAATTCGAAGCAGCAAC	1440
Q	y		1441	CAATGACTCTGCTGCTTGTGATGTCGCCCCACCCCAAGGCGCTCAAGACAAATATGCTATGG	1500
D	b		1441	CAATGACTCTGCTGCTTGTGATGTCGCCCCACCCCAAGGCGCTCAAGACAAATATGCTATGG	1500
Q	y		1501	CGGGAAGACCCGAGGGGAGAGATGAGAGAACCCGAGAACAGATGGCTTCTGAAGTACACCA	1560
D	b		1501	CGGGAAGACCCGAGGGGAGAGATGAGAGAACCCGAGAACAGATGGCTTCTGAAGTACACCA	1560
Q	y		1561	CAACAAGGGCATCTGGAAGACCGCTGTTTGTCCGTGGAAGAAAGAACCTGTGTCTT	1620
D	b		1561	CAACAAGGGCATCTGGAAGACCGCTGTTTGTCCGTGGAAGAAAGAACCTGTGTCTT	1620
Q	y		1621	CCACCCCTCTTTGAGGGTGGGCTCTGTACAGATTGCCGGATGCTTCTAGAGCTCTT	1680
D	b		1621	CCACCCCTCTTTGAGGGTGGGCTCTGTACAGATTGCCGGATGCTTCTAGAGCTCTT	1680

[illegible]

OY	2761	CCGTCAGAAC	CTCTG	GGGAGG	TCCTGG	AGTGTAC	CCGGCAT	CCAGAC	CTGTTT	CCCC	2825			
Dp	2701	CCGTCAGAAC	CTCTG	GGGAGG	TCCTGG	AGTGTAC	CCGGCAT	CCAGAC	CTGTTT	CCCC	2760			
OY	2821	CTTGAAAG	CACTACT	TTGGCT	GGAAAT	CTTAC	CCAGAC	CTGGGG	AGCTTC	GGGACGA	2880			
Dp	2761	CTTGAAAG	CACTACT	TTGGCT	GGAAAT	CTTAC	CCAGAC	CTGGGG	AGCTTC	GGGACGA	2820			
OY	2881	GCACAGT	CCAGAGT	ACCCCT	CCCTGA	AGGAC	CTCAC	CTGT	CCCC	TTTATAGCTAC	2940			
Dp	2821	GCACAGT	CCAGAGT	ACCCCT	CCCTGA	AGGAC	CTCAC	CTGT	CCCC	TTTATAGCTAC	2880			
OY	2941	TGTGTGGG	CCCTCAC	ATCAG	CTGAC	CTTGTCT	CTGT	CAGTGG	AGCAG	AGCTC	3000			
Dp	2881	TGTGTGGG	CCCTCAC	ATCAG	CTGAC	CTTGTCT	CTGT	CAGTGG	AGCAG	AGCTC	2940			
OY	3001	CTGGCC	CTTTCAG	GGGAG	CCCCGG	TGCTCC	CTCC	CTGTGC	ACAC	CTAGAC	CTGTGCT	3060		
Dp	2941	CTGGCC	CTTTCAG	GGGAG	CCCCGG	TGCTCC	CTCC	CTGTGC	ACAC	CTAGAC	CTGTGCT	3000		
OY	3061	TAGAGT	AGCCCCG	ATG	GTGTCAT	GTTCCTTAC	CCCTGA	AACTTTAA	ACTTGA	AGTAG	3120			
Dp	3001	TAGAGT	AGCCCCG	ATG	GTGTCAT	GTTCCTTAC	CCCTGA	AACTTTAA	ACTTGA	AGTAG	3060			
OY	3121	TGAGTAG	ATG	GGTTTCTTTT	ACCCTCTAG	TTTTC	ACTC	ACAC	AGTATG	GGCTAAGT	3180			
Dp	3061	TGAGTAG	ATG	GGTTTCTTTT	ACCCTCTAG	TTTTC	ACTC	ACAC	AGTATG	GGCTAAGT	3120			
OY	3181	ACGAAAAA	CAACAA	CAAAAAAC	AGAAAC	CAAAAAA	CAAAAAA	CAAAAAA	CAAAAAA	CAAAAAA	CTCAAGCTGCTTA	3240		
Dp	3121	ACGAAAAA	CAACAA	CAAAAAAC	AGAAAC	CAAAAAA	CAAAAAA	CAAAAAA	CAAAAAA	CAAAAAA	CTCAAGCTGCTTA	3180		
OY	3241	GATCA	AGGTTAC	TGCTG	CAAAAT	CACTAG	AGATTTT	GTTTTAA	GTAA	AGAC	CCGTGCTCA	3300		
Dp	3181	GATCA	AGGTTAC	TGCTG	CAAAAT	CACTAG	AGATTTT	GTTTTAA	GTAA	AGAC	CCGTGCTCA	3240		
OY	3301	CATT	TGCTG	AGAGAT	GCATTT	GTGTA	ATGTGG	GTCA	AGAT	GAAGCA	AGGTCAAGGGGCGAA	3360		
Dp	3241	CATT	TGCTG	AGAGAT	GCATTT	GTGTA	ATGTGG	GTCA	AGAT	GAAGCA	AGGTCAAGGGGCGAA	3300		
OY	3361	AAAAA	TTCC	CCCCCT	CTCCCC	CCAG	AGATTTTGA	AGTGTAT	GTGTTA	AGTCTTCC	3420			
Dp	3301	AAAAA	TTCC	CCCCCT	CTCCCC	CCAG	AGATTTTGA	AGTGTAT	GTGTTA	AGTCTTCC	3360			
OY	3421	TGGC	ACCTTCCC	CTTGTGG	TATG	CAAGGG	CTGAAG	CTCTGT	GTGCTT	GTAG	CAATTTCC	3480		
Dp	3361	TGGC	ACCTTCCC	CTTGTGG	TATG	CAAGGG	CTGAAG	CTCTGT	GTGCTT	GTAG	CAATTTCC	3420		
OY	3481	CAGAT	ANTANT	ATCAG	AGGATG	ACATAC	ACCAC	CTTTTAA	GGGCTTTT	CCCTGG	AGGGG	3540		
Dp	3421	CAGAT	ANTANT	ATCAG	AGGATG	ACATAC	ACCAC	CTTTTAA	GGGCTTTT	CCCTGG	AGGGG	3480		
OY	3541	CCCAT	TGGGCT	TA	CTCCTC	ACGAA	AC	TG	AGTGTGA	TGTTG	AGAGCTGGG	3600		
Dp	3481	CCCAT	TGGGCT	TA	CTCCTC	ACGAA	AC	TG	AGTGTGA	TGTTG	AGAGCTGGG	3540		
OY	3601	TGGAG	TGGCC	CTTTCC	AGAG	TGTAG	AGGAT	TAC	CAAG	AGAGAA	AGCTTACCATTC	3660		
Dp	3541	TGGAG	TGGCC	CTTTCC	AGAG	TGTAG	AGGAT	TAC	CAAG	AGAGAA	AGCTTACCATTC	3600		
OY	3661	CCCA	CTCCCTT	TG	CCCAAT	AGAGGG	CCCA	AGTCCCA	AC	CTAG	GTCCCAAGAACCC	3720		
Dp	3601	CCCA	CTCCCTT	TG	CCCAAT	AGAGGG	CCCA	AGTCCCA	AC	CTAG	GTCCCAAGAACCC	3660		
OY	3721	CTAG	CTTCC	TATG	AGAG	CTAG	ACCA	AGAC	CAT	CGTTC	CTTATCTG	AGAGCTGT	3780	
Dp	3661	CTAG	CTTCC	TATG	AGAG	CTAG	ACCA	AGAC	CAT	CGTTC	CTTATCTG	AGAGCTGT	3720	
OY	3781	TGGG	GAAC	TA	CTAG	AAAA	CTTCTG	GAGAT	GTTTTAA	AGCTTTT	TACCC	ACAGAT	AGAT	3840
Dp	3721	TGGG	GAAC	TA	CTAG	AAAA	CTTCTG	GAGAT	GTTTTAA	AGCTTTT	TACCC	ACAGAT	AGAT	3780

OY		3841	GTTGTTTTAAAGGGGTGCTTTTATTTTAAAGGCATCACTGACGAATTAGAAGCGCATTTTCAG	3900
Dd		3781	GTGTTTTTAAGGGGTGCTTTTATTTTAAAGGCATCACTGACGAATTAGAAGCGCATTTTCAG	3840
OY		3901	AAATGCCATCGTAATGGTTTTTAAACACCTTTTACCATAATTACAGTGCTATTATTATAGA	3960
Dd		3841	AAATGCCATCGTAATGGTTTTTAAACACCTTTTACCATAATTACAGTGCTATTATTATAGA	3900
OY		3961	AGCAGACAACACTTCCTTTTATGACCTCCAGACTTCATATTTTCATGTTACCATTATTTTTT	4020
Dd		3901	AGCAGACAACACTTCCTTTTATGACCTCCAGACTTCATATTTTCATGTTACCATTATTTTTT	3960
OY		4021	GTAACCTGCAAGGTGTGGGCTTTTGTACTTCAAGGTGTGGGAGAGACTGCTGTTT	4080
Dd		3961	GTAACCTGCAAGGTGTGTGGGCTTTTGTACTTCAAGGTGTGGGAGAGACTGCTGTTT	4020
OY		4141	TTACCTTTCTGCTCACCTGCTTTCTAATTTTTAGGTCAAAGATGACAGATGCCCAAGT	4140
Dd		4081	TTACCTTTCTGCTCACCTGCTTTCTAATTTTTAGGTCAAAGATGACAGATGCCCAAGT	4080
RESULT 4		AFI51970	4278 bp	mRNA linear ROD 08-JUN-2000
LOCUS		AFI51970	Mus musculus DNA cytosine-specific methyltransferase isoform 2	(Dnmt3b) mRNA, complete cds.
DEFINITION		AFI51970	AFI51970.1 GI:8347119	
ACCESSION		AFI51970		
VERSION		AFI51970.1	GI:8347119	
KEYWORDS				
SOURCE				
ORGANISM				Mus musculus.
REFERENCE				Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS				Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE				Yin,B., Chen,Y.T., Zhu,M., Luo,Y.J., Zhu,N., Xu,S.C., Wu,G.Y. and Shen,Y.
JOURNAL				Cloning of full-length Dnmt3b cDNA and its alternative splicing
REFERENCE				isoforms in mouse embryonic tissue
AUTHORS				Unpublished
TITLE				2 (bases 1 to 4278)
JOURNAL				Submitted (17-MAY-1999) Department of Biochemistry and Molecular
REFERENCE				Biology, Institution of Basic Medical Sciences, Chinese Academy of
AUTHORS				Medical Sciences, Dong Dan Shan Tiao 5, Beijing City 100005,
TITLE				P.R.China
JOURNAL				Direct Submission
REFERENCE				Submitted (17-MAY-1999) Department of Biochemistry and Molecular
AUTHORS				Biology, Institution of Basic Medical Sciences, Chinese Academy of
TITLE				Medical Sciences, Dong Dan Shan Tiao 5, Beijing City 100005,
JOURNAL				P.R.China
REFERENCE				Submitted (17-MAY-1999) Department of Biochemistry and Molecular
AUTHORS				Biology, Institution of Basic Medical Sciences, Chinese Academy of
TITLE				Medical Sciences, Dong Dan Shan Tiao 5, Beijing City 100005,
JOURNAL				P.R.China
REFERENCE				Submitted (17-MAY-1999) Department of Biochemistry and Molecular
AUTHORS				Biology, Institution of Basic Medical Sciences, Chinese Academy of
TITLE				Medical Sciences, Dong Dan Shan Tiao 5, Beijing City 100005,
JOURNAL				P.R.China
REFERENCE				Submitted (17-MAY-1999) Department of Biochemistry and Molecular
AUTHORS				Biology, Institution of Basic Medical Sciences, Chinese Academy of
TITLE				Medical Sciences, Dong Dan Shan Tiao 5, Beijing City 100005,
JOURNAL				P.R.China
REFERENCE				Submitted (17-MAY-1999) Department of Biochemistry and Molecular
AUTHORS				Biology, Institution of Basic Medical Sciences, Chinese Academy of
TITLE				Medical Sciences, Dong Dan Shan Tiao 5, Beijing City 100005,
JOURNAL				P.R.China
REFERENCE				Submitted (17-MAY-1999) Department of Biochemistry and Molecular
AUTHORS				Biology, Institution of Basic Medical Sciences, Chinese Academy of
TITLE				Medical Sciences, Dong Dan Shan Tiao 5, Beijing City 100005,
JOURNAL				P.R.China
REFERENCE				Submitted (17-MAY-1999) Department of Biochemistry and Molecular
AUTHORS				Biology, Institution of Basic Medical Sciences, Chinese Academy of
TITLE				Medical Sciences, Dong Dan Shan Tiao 5, Beijing City 100005,
JOURNAL				P.R.China
REFERENCE				Submitted (17-MAY-1999) Department of Biochemistry and Molecular
AUTHORS				Biology, Institution of Basic Medical Sciences, Chinese Academy of
TITLE				Medical Sciences, Dong Dan Shan Tiao 5, Beijing City 100005,
JOURNAL				P.R.China
REFERENCE				Submitted (17-MAY-1999) Department of Biochemistry and Molecular
AUTHORS				Biology, Institution of Basic Medical Sciences, Chinese Academy of
TITLE				Medical Sciences, Dong Dan Shan Tiao 5, Beijing City 100005,
JOURNAL				P.R.China
REFERENCE				Submitted (17-MAY-1999) Department of Biochemistry and Molecular
AUTHORS				Biology, Institution of Basic Medical Sciences, Chinese Academy of
TITLE				Medical Sciences, Dong Dan Shan Tiao 5, Beijing City 100005,
JOURNAL				P.R.China
REFERENCE				Submitted (17-MAY-1999) Department of Biochemistry and Molecular
AUTHORS				Biology, Institution of Basic Medical Sciences, Chinese Academy of
TITLE				Medical Sciences, Dong Dan Shan Tiao 5, Beijing City 100005,
JOURNAL				P.R.China
REFERENCE				Submitted (17-MAY-1999) Department of Biochemistry and Molecular
AUTHORS				Biology, Institution of Basic Medical Sciences, Chinese Academy of
TITLE				Medical Sciences, Dong Dan Shan Tiao 5, Beijing City 100005,
JOURNAL				P.R.China
REFERENCE				Submitted (17-MAY-1999) Department of Biochemistry and Molecular
AUTHORS				Biology, Institution of Basic Medical Sciences, Chinese Academy of
TITLE				Medical Sciences, Dong Dan Shan Tiao 5, Beijing City 100005,
JOURNAL				P.R.China
REFERENCE				Submitted (17-MAY-1999) Department of Biochemistry and Molecular
AUTHORS				Biology, Institution of Basic Medical Sciences, Chinese Academy of
TITLE				Medical Sciences, Dong Dan Shan Tiao 5, Beijing City 100005,
JOURNAL				P.R.China
REFERENCE				Submitted (17-MAY-1999) Department of Biochemistry and Molecular
AUTHORS				Biology, Institution of Basic Medical Sciences, Chinese Academy of
TITLE				Medical Sciences, Dong Dan Shan Tiao 5, Beijing City 100005,
JOURNAL				P.R.China
REFERENCE				Submitted (17-MAY-1999) Department of Biochemistry and Molecular
AUTHORS				Biology, Institution of Basic Medical Sciences, Chinese Academy of

Query Match	Best Local Similarity	Score	DB 10:	Length	4278:
Matches	Conservative	Mismatches	Indels	Gaps	
10	96.6%	98.5%	0	2	60
153	90.0%	90.0%	0	2	60
70	90.0%	90.0%	0	2	60
213	90.0%	90.0%	0	2	60
130	90.0%	90.0%	0	2	60
273	90.0%	90.0%	0	2	60
190	90.0%	90.0%	0	2	60
333	90.0%	90.0%	0	2	60
250	90.0%	90.0%	0	2	60
393	90.0%	90.0%	0	2	60
310	90.0%	90.0%	0	2	60
453	90.0%	90.0%	0	2	60
370	90.0%	90.0%	0	2	60
513	90.0%	90.0%	0	2	60
430	90.0%	90.0%	0	2	60
573	90.0%	90.0%	0	2	60
490	90.0%	90.0%	0	2	60
633	90.0%	90.0%	0	2	60
550	90.0%	90.0%	0	2	60
693	90.0%	90.0%	0	2	60
610	90.0%	90.0%	0	2	60
753	90.0%	90.0%	0	2	60
670	90.0%	90.0%	0	2	60
813	90.0%	90.0%	0	2	60
730	90.0%	90.0%	0	2	60
873	90.0%	90.0%	0	2	60
790	90.0%	90.0%	0	2	60
933	90.0%	90.0%	0	2	60

QY	85	ATCAGTTGACTGTAGCAGACATGCGAGATCAAGAGGGGTATGATATACACACAGGTGATGC	909
Db	999	ATCAGTTGACTGTAGCAGACATGCGAGATCAAGAGGGGTATGATATACACACAGGTGATGC	1055
QY	910	AGAGGCGAGATATGAGAGACAGACAGATATCAAGATATGATATGAAAGTTGGATTAGTGA	969
Db	1053	AGAGGCGATATATGAGAGACAGACAGATATCAAGATATGATATGAAAGTTGGATTAGTGA	1112
QY	970	CCTGCTGGGGAAAGATCAAGGGCTTCTCTGCTGGCTGCTGCATGGTGGTCTCTCGAA	1028
Db	1113	CCTGCTGGGGAAAGATCAAGGGCTTCTCTGCTGGCTGCTGCATGGTGGTCTCTCGAA	1177
QY	1030	AGCCACCTCCAGACGACAGCGCATGCCCGAATGCGCTGGGTACAGTGGTTGGTATGG	1089
Db	1173	AGCCACCTCCAGACGACGACCGCATGCCCGAATGCGCTGGGTACAGTGGTTGGTATGG	1232
QY	1090	CAAGTTTTCGTAGATCTGTGCTGACAAACTGGTGGCTGGGGGCTTCAGCCAGCACTT	1149
Db	1233	CAAGTTTTCGTAGATCTGTGCTGACAAACTGGTGGCTGGGGGCTTCAGCCAGCACTT	1292
QY	1150	TAACTGTGCTACCTTCAATTAAGCTGGTTCTTATAGGAAGGCCATGTACACACTGTGA	1209
Db	1293	TAACTGTGCTACCTTCAATTAAGCTGGTTCTTATAGGAAGGCCATGTACACACTGTGA	1355
QY	1210	GAAAGCCAGAGGTTGAGAGTGGCAGACCTTCTCAGCACTCTCTGAGAGTCACTGAGGA	1269
Db	1353	GAAAGCCAGAGGTTGAGAGTGGCAGACCTTCTCAGCACTCTCTGAGAGTCACTGAGGA	1412
QY	1270	CCAGCTGAAGCCCATCTGAGTGGGGCCACAGGTGGCTTCAAGGCTACTGGATTCAGAGG	1328
Db	1413	CCAGCTGAAGCCCATCTGAGTGGGGCCACAGGTGGCTTCAAGGCTACTGGATTCAGAGG	1472
QY	1330	CCTCAAAACCCACAAACAAACCAACCAAGTGGTTATATAGTGAAAGGTGGTGTCAACAG	1389
Db	1473	CCTCAAAACCCACAAACAAACCAACCAAGTGGTTATATAGTGAAAGGTGGTGTCAACAG	1497
QY	1390	TAGAAACTTAGAACCCAGAGACCGGAGAACCAATGTGAGAGGCGCAACCAATATCTC	1449
Db	1498	-----GAGAACCAATGTGAGAGGCGCAACCAATATCTC	1533
QY	1450	TGCTGCTTTCGATCCCCCCCCCAACCCAGCGCTCAAGACAAATATGCTATGGCCGAGGA	1509
Db	1533	TGCTGCTTTCGATCCCCCCCCCAACCCAGCGCTCAAGACAAATATGCTATGGCCGAGGA	1592
QY	1510	CCGAGGGAGATATGAGAGAGCCGAGAACGATGGCTTCTGAGTCAACAAACAAGGG	1569
Db	1593	CCGAGGGAGATATGAGAGAGCCGAGAACGATGGCTTCTGAGTCAACAAACAAGGG	1652
QY	1570	CAATCTGGAAGACCGCTGTTTGTCTCTGTGGAAGAAACCCGTGTCTTCCACCCCT	1628
Db	1653	CAATCTGGAAGACCGCTGTTTGTCTCTGTGGAAGAAACCCGTGTCTTCCACCCCT	1712
QY	1630	CTTTTGAGGGTGGGCTCTGTCAAGTATGCGCGGATTCGCTTCAGAGTCTTTCACTATGA	1689
Db	1713	CTTTTGAGGGTGGGCTCTGTCAAGTATGCGCGGATTCGCTTCAGAGTCTTTCACTATGA	1772
QY	1690	TGATGAGAGCGGTATCACTCCTACTGACACGATGCTGCTGAGGGCGGTAAATGCTGCT	1749
Db	1773	TGATGAGAGCGGTATCACTCCTACTGACACGATGCTGCTGAGGGCGGTAAATGCTGCT	1832
QY	1750	GTGCGATTAACAACAGTGTGCTGAGAGTGTCTGTGTGAGATGTCTGAGAGGTGCTGGTGGG	1809
Db	1833	GTGCGATTAACAACAGTGTGCTGAGAGTGTCTGTGTGAGATGTCTGAGAGGTGCTGGTGGG	1892
QY	1810	CGCAGGCGACAGCTGAGAGATGCCAAGCTGCAAGAAACCTTGGAGTCTATATATGCTGCC	1869
Db	1893	CGCAGGCGACAGCTGAGAGATGCCAAGCTGCAAGAAACCTTGGAGTCTATATATGCTGCC	1952
QY	1870	TCAGGCGTCCATGGGGTCTCTCCAGCAGAGAAAGATTGGAACATGCCCTGCACAGCTT	1929
Db	1953	TCAGGCGTCCATGGGGTCTCTCCAGCAGAGAAAGATTGGAACATGCCCTGCACAGCTT	2012
QY	1930	CTTCACTACTGATCTCGACTGGAAGAAATTTGAGCCACCAAGTGTATCCACCAATTCC	1989

Db	2013	CTTCCATACGATCCCGACCTGGAAAGATTGGACCAACCAAGTTGATACCCACCAATTCC	2072
Oy	1990	TGCAGCCAAAAGAGGCCCATTAAGACGTCGTCGTGTTGATGGAAATTGCACAGGGGTA	2049
Db	2073	TGCACCCAAAAGAGGCCCATTAAGAGTCGTCTCTGTTATGCAATTGCACAGGGGTA	2132
Oy	2050	CTTGGTCTCAAGGAGTTGGGTATTTAAATGGAAAATACATTCGCTCCGAGTCTGTGC	2109
Db	2133	CTTGGTCTCAAGGAGTTGGGTATTTAAATGGAAAATACATTCGCTCCGAGATCTGTGC	2192
Oy	2110	AGAGTCATCGCTGTGGGAACGTGTTAAGCATGGAAGCCAGATCAAAATATGTCAATGACGT	2169
Db	2193	AGAGTCATCGCTGTGGGAACGTGTTAAGCATGGAAGCCAGATCAAAATATGTCAATGACGT	2252
Oy	2170	CCGGAANAATCACCAAGAAAATATTGGAAGATGGGGCCGTTGACATTTGGTGAATGTGTG	2229
Db	2253	CCGGAANAATCACCAAGAAAATATTGGAAGATGGGGCCGTTGACATTTGGTGAATGTGTG	2312
Oy	2230	AAGCCATGCAATGATCTCTCTAACGTCATCCGTGCCGCAAAAGTTTATATAGGGCAC	2289
Db	2313	AAGCCATGCAATGATCTCTCTAACGTCATCCGTGCCGCAAAAGTTTATATAGGGCAC	2372
Oy	2290	AGGAAGGCTCTTCTTGATTTACCACTTGCTGAATTATACCCGCCCAAGAGGGGCA	2349
Db	2373	AGGAAGGCTCTTCTTGATTTACCACTTGCTGAATTATACCCGCCCAAGAGGGGCA	2432
Oy	2350	CACCGTCACATTTCTTGATTTGTCGAGAAATGTTGGCCATGAAATGAAATACACAA	2409
Db	2433	CACCGTCACATTTCTTGATTTGTCGAGAAATGTTGGCCATGAAATGAAATACACAA	2492
Oy	2410	AGACATCTCAAGATTCCTGGGACGTGTAACCCAGATGATCGATGCATCAAGGTGTCTGC	2469
Db	2493	AGACATCTCAAGATTCCTGGGACGTGTAACCCAGATGATCGATGCATCAAGGTGTCTGC	2552
Oy	2470	TGCTCACAGGGCCCGGACTTCTGGGGTAACCTACCCGAATGAACAGCGCGTGAATGC	2529
Db	2553	TGCTCACAGGGCCCGGACTTCTGGGGTAACCTACCCGAATGAACAGCGCGTGAATGC	2612
Oy	2530	TTCAAAAGATGATAGCTCGAGCTGCACAGACTCCCTCGAGATTCAAGTGGACACCAAGTT	2589
Db	2613	TTCAAAAGATGATAGCTCGAGCTGCACAGACTCCCTCGAGATTCAAGTGGACACCAAGTT	2672
Oy	2590	AAAGAAAGTGCAGACAATPACCAACCAAGTCGAATCTCATCAGACAGGGCAAAACCACT	2649
Db	2673	AAAGAAAGTGCAGACAATPACCAACCAAGTCGAATCTCATCAGACAGGGCAAAACCACT	2732
Oy	2650	TTTCCCTGTAGTCATGAATGCGAAGACGACGTTTTGTGTGTCACCTGAGCTCGAAAGAT	2709
Db	2733	TTTCCCTGTAGTCATGAATGCGAAGACGACGTTTTGTGTGTCACCTGAGCTCGAAAGAT	2792
Oy	2710	CTTGGGCTTCCCTGCTCATACAGGACGTCGTCCAACATGGGCGCGGCCCCCTGCAGAA	2769
Db	2793	CTTGGGCTTCCCTGCTCATACAGGACGTCGTCCAACATGGGCGCGGCCCCCTGCAGAA	2852
Oy	2770	GCTGCTGGGAGGTCCTGAGAGTACCGGTCATCAGACACCTGTTCCCTCTTGAAAGTA	2839
Db	2853	GCTGCTGGGAGGTCCTGAGAGTACCGGTCATCAGACACCTGTTCCCTCTTGAAAGTA	2912
Oy	2830	CTACTTTGACCTGGAAATAGTCTACCCAGAGACGGGGAAGCTCGGTGCACAGCAAGTGC	2889
Db	2913	CTACTTTGACCTGGAAATAGTCTACCCAGAGACGGGGAAGCTCGGTGCACAGCAAGTGC	2972
Oy	2890	CAGAGTCACCCCTCCCTGAAGGACCTCACCTGTCCCTTTTAGTCACACGTGTGGGG	2949
Db	2973	CAGAGTCACCCCTCCCTGAAGGACCTCACCTGTCCCTTTTAGTCACACGTGTGGGG	3032
Oy	2950	CCTCACATCACTGTACCTGACGCTTCTCTGCTCAGGTGGAGAGACGCTCCTGGCCTT	3009
Db	3033	CCTCACATCACTGTACCTGACGCTTCTCTGCTCAGGTGGAGAGACGCTCCTGGCCTT	3092
Oy	3010	GCAGGGAGGCCCGGAGTCCCTCCGTCGTCACACCTCAGACCTGGCTGCTTATAGATAGC	3069

```

Db      3093 GCAGGGAGCCCGGTCGCCCTCCGTCGTCGACAGCTCAGACCTGGCTTGTAGAGTAC 3152
Oy      3070 CCGGCATGTCGTCATGTCCTTACCTGAACTTTAAACTGTAAGTAGTAGTAGA 3129
Db      3153 CCGGCATGTCGTCATGTCCTTACCTGAACTTTAAACTGTAAGTAGTAGA 3212
Oy      3130 TGGCTTCTTTTACCTCTGAGTTTATCACTCAGAAAGTAGTAGTAGA 3189
Db      3213 TGGCTTCTTTTACCTCTGAGTTTATCACTCAGAAAGTAGTAGTAGA 3272
Oy      3190 ACAACAAAAACAGAAACAAAAAACAACAACTCAGAGCTCTTGTAGTAG 3249
Db      3273 ACAACAAAAACAGAAACAAAAAACAACAACTCAGAGCTCTTGTAGTAG 3332
Oy      3250 TTCAATGTCGAAATTCATGATTTGTTTAAAGTAAACCCGCTCCACATTTGCTG 3309
Db      3333 TTCAATGTCGAAATTCATGATTTGTTTAAAGTAAACCCGCTCCACATTTGCTG 3392
Oy      3310 GAGGATGCTATTGTAATGTCGGCTCAGATGAGCAAGTCAAGGGGCCAAAAAATTC 3369
Db      3393 GAGGATGCTATTGTAATGTCGGCTCAGATGAGCAAGTCAAGGGGCCAAAAAATTC 3452
Oy      3370 CCGCTCTCCCGCAGAGTATTTGAAGATGATGTTTAAAGTAAAGTCTTCTGCACT 3429
Db      3453 CCGCTCTCCCGCAGAGTATTTGAAGATGATGTTTAAAGTAAAGTCTTCTGCACT 3512
Oy      3430 CCGCTCTCTGTTGTAACAAGGGCTGAAGTCTGTTGCTTGTAGATTTCCAGATAT 3489
Db      3513 CCGCTCTCTGTTGTAACAAGGGCTGAAGTCTGTTGCTTGTAGATTTCCAGATAT 3572
Oy      3490 GATGTCAGCAGGATGATCATCACCACCTTTAGGGCTTTCCCTGGCAGGGCCCAT 3549
Db      3573 GATGTCAGCAGGATGATCATCACCACCTTTAGGGCTTTCCCTGGCAGGGCCCAT 3632
Oy      3550 CTAGTCTCTCAGAACTGCAATGATGTTTGAAGTCTCAGAGAGGTGGTGGAGTGG 3609
Db      3633 CTAGTCTCTCAGAACTGCAATGATGTTTGAAGTCTCAGAGAGGTGGTGGAGTGG 3692
Oy      3610 CCTCTCCAGGTGAGGATACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3669
Db      3693 CCTCTCCAGGTGAGGATACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3752
Oy      3670 TCTTGGCCAAATGAGGGGCCAGTCCCAACAGCTCAGAGTCCCAACAGTCCCAAG 3729
Db      3753 TCTTGGCCAAATGAGGGGCCAGTCCCAACAGCTCAGAGTCCCAACAGTCCCAAG 3812
Oy      3730 CATGAGAACTGAGACCAAGAACATGCTCCCTTATCTGAGCAGTGTGGGGA 3789
Db      3813 CATGAGAACTGAGACCAAGAACATGCTCCCTTATCTGAGCAGTGTGGGGA 3872
Oy      3790 ACAGTGAACCTCTCTGAGATGTTTAAAGTCTTTTACCCAGATAGATGTGTTT 3849
Db      3873 ACAGTGAACCTCTCTGAGATGTTTAAAGTCTTTTACCCAGATAGATGTGTTT 3932
Oy      3850 AGGGGTGCTTTTATAGGGGATCAGTGAATGAAGAGTGCATTTCAAGAAATGC 3909
Db      3933 AGGGGTGCTTTTATAGGGGATCAGTGAATGAAGAGTGCATTTCAAGAAATGC 3992
Oy      3910 CGTAATGCTTTTAAACACCTTTTACCTAATTAACAGTGTCTTATGAAGCAG 3969
Db      3993 CGTAATGCTTTTAAACACCTTTTACCTAATTAACAGTGTCTTATGAAGCAG 4052
Oy      3970 CAGTCTTTTATAGCTCTCAGACTCTTATTTTACCTTATGATGATGATGATG 4029
Db      4053 CAGTCTTTTATAGCTCTCAGACTCTTATTTTACCTTATGATGATGATGATG 4112
Oy      4030 AAGGTGGGCTTTTGTACTTCAAGGTGTGGGAGAGAGTGCCTTGTTCACAGTT 4089
Db      4113 AAGGTGGGCTTTTGTACTTCAAGGTGTGGGAGAGAGTGCCTTGTTCACAGTT 4172
Oy      4090 GTCTCCACTGTTCTTAATTTTATAGTGCAGAAAGATGACAGATGCCAGAGTTT 4149
Db      4173 GTCTCCACTGTTCTTAATTTTATAGTGCAGAAAGATGACAGATGCCAGAGTTT 4232

```

```

Oy      4150 TGGTGAATTAAGTTGATTTCTCTPAAAAAAGTAAAAAAGTAAAAAAGTAAAAA 4195
Db      4233 TGGTGAATTAAGTTGATTTCTCTPAAAAAAGTAAAAAAGTAAAAAAGTAAAAA 4278

RESULT 5
AF151973
LOCUS
DEFINITION
Mus musculus DNA cytosine-specific methyltransferase isoform 5
(Dmmt3b) mRNA, complete cds.
ACCESSION
AF151973
VERSION
AF151973.1 GI:8347127
KEYWORDS
SOURCE
Mus musculus.
MUS MUSCULUS.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 4223)
Yip, B., Chen, Y.T., Zhu, M., Luo, Y.J., Zhu, N., Xu, S.C., Wu, G.Y. and
Shen, Y.
Cloning of full-length Dmmt3b cDNA and its alternative splicing
isoforms in mouse embryonic tissue
Unpublished
2 (bases 1 to 4223)
Yip, B., Chen, Y.T., Zhu, M., Luo, Y.J., Zhu, N., Xu, S.C., Wu, G.Y. and
Shen, Y.
Direct Submission
Submitted (17-May-1999) Department of Biochemistry and Molecular
Biology, Institution of Basic Medical Sciences, Chinese Academy of
Medical Sciences, Dong Dan San Tiao 5, Beijing City 100005,
P.R.China

FEATURES
Source
location/Qualifiers
1..4223
/organism="Mus musculus"
/strain="KM"
/db_xref="taxon:10090"
/dev_stage="8-9 day old embryo"
1..4223
/gene="Dmmt3b"
297..2876
/gene="Dmmt3b"
/function="methylates cytosine in DNA"
/note="alternatively spliced"
/codon_start=1
/product="DNA cytosine-specific methyltransferase isoform
5"
/protein_id="AAF74519.1"
/db_xref="GI:8347128"
/translation="MKGDSRLNREERASGYEPTIVNGNFSDSDPKADAPPEVLE
ALTEPVCTPEPTRGRSSRLSKREVSSLLATYQDMTGDGRDDEVDNGSDILMPK
LTREKDRTRSESFAVTRHNSGTSLSERKAPSRITRGQGRHVEYVEFPATR
SRRRASASASTPWSVSDPMEVEYTRSTVSTPSVDLSQDDEGMDTTOVDASISY
GDSVEYODDKREFGIDLVWGRIKGSMMPAVSVSKATSKQAMGVOMVGDKGK
SEISDAKLVALGFLFSOHNLATFNKLIVYRRAMVYTLKARVAGKTPSSPGSELED
OLKPLMEAHAGHGFPTGTEGLKPKNKQVYVSKVRSRSDSHNLEPRREKNSRRRTN
DSASESPPRRLKTNISYGGDRDESDSRPMASEYVNNKNGNLEDRLSGCKRPVS
FHLPEGLCOCDRDLFTYMDDEDISTCYVCCGRELILCSNISCRCRCVBC
LELVAGATADAKIQEPWSCYMLPQRCHEVLRKRMNRLDDEFTTDPDLFEFEP
PKLPAIPAAKRRPRLVSLFDGATGLVKEIGIKVEKYIASVCAESIAVGKVA
EGQIKYVNDVRKITKKNLEWGPEDLVIGSPKNDLSVNPARKGLVGTSLFEFEPY
HLNLTREKEDNREPMFENYVAMKVNDRKDISRLFACNPMIDATVGNASHARY
FMGNLPGNNRPVMSKNDKLLQDCLSEPRAKLAKKYOTITTKSMISIQGKNQLEPVY
MKGDDVLMCTELEIRIFGFPAAHYDVSNMKGARKOKLLGRSMVSVIINLAPLKOYF
ACE"

BASE COUNT
1051 a 1084 c 1147 g 941 t

ORIGIN
Query Match 94.0%; Score 3942.8; DB 10; Length 4223;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3938; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Oy      251 GCCTTCTTTCAGGAACAATGAGGAGACAGACAGACATCTGATGATGAAGAGAGGTGCC 310

```

Db 279 GCCTCAGCAGAGAAACAAATGAAAGGAGAGACATCTGAAATGAAAGAGGGGCGC 338
OY 311 AGCGGATATGAGAGATGATTATGTTAATGGAACTTCAGTACCAAGTCCTGAGACAG 370
Db 339 AGCGGATATGAGAGATGATTATGTTAATGGAACTTCAGTACCAAGTCCTGAGACAG 398
OY 371 AAGATGCTCCCTCACCCTCAGTCTTGGAGGCAATGTGCACAGACCAGTGTGCACACA 430
Db 399 AAGATGCTCCCTCACCCTCAGTCTTGGAGGCAATGTGCACAGACCAGTGTGCACACA 458
OY 431 GAGACAGAGGCGGAGGTCAAGCTCCGCTCTCTAAGAGGAGAGTCTCAGCTCTCTG 490
Db 459 GAGACAGAGGCGGAGGTCAAGCTCCGCTCTCTAAGAGGAGAGTCTCAGCTCTCTG 518
OY 491 AATTACAGCAGATACAGAGAGATGAGACAGAGATGATGATGATGATGATGATGAT 550
Db 519 AATTACAGCAGATACAGAGAGATGAGACAGAGATGATGATGATGATGATGATGAT 578
OY 551 GCGCTGATATCTTAAATGCAAAAGCTCAACCTGAGACCAAGACACAGACCGCTCT 610
Db 579 GCGCTGATATCTTAAATGCAAAAGCTCAACCTGAGAACCAAGACACAGACCGCTCT 638
OY 611 GAAAGCGCGGCTGTCCGAAACCCGACATAGCAATGAGACCTCCAGCTTGGAGGCAAGA 670
Db 639 GAAAGCGCGGCTGTCCGAAACCCGACATAGCAATGAGACCTCCAGCTTGGAGGCAAGA 698
OY 671 GCCTCCCGCAGAAATCACCCGAGGTGCGACAGGCGCCACCATGTGCAGAGTACCTGTG 730
Db 699 GCCTCCCGCAGAAATCACCCGAGGTGCGACAGGCGCCACCATGTGCAGAGTACCTGTG 758
OY 731 GAGTTTCGGCTACAGGTCTCGGAGAGCTCGAGCATCTCTTCAGCAAGCACCCATGG 790
Db 759 GAGTTTCGGCTACAGGTCTCGGAGAGCTCGAGCATCTCTTCAGCAAGCACCCATGG 818
OY 791 TCATCCCTCGCGAGGTGAGCTTCATGAAAGAGTGCACCTCAAGAGCGTCAGTACCCA 850
Db 819 TCATCCCTCGCGAGGTGAGCTTCATGAAAGAGTGCACCTCAAGAGCGTCAGTACCCA 878
OY 851 TCAGTTGACTGAGCGCAGGATGATGAAAGAGTGCACCTCAAGAGCGTCAGTACCCA 910
Db 879 TCAGTTGACTGAGCGCAGGATGATGAAAGAGTGCACCTCAAGAGCGTCAGTACCCA 938
OY 911 GAGAGCAGATGAGAGACAGACAGATGATGAAAGAGTGCACCTCAAGAGCGTCAGTAC 970
Db 939 GAGAGCAGATGAGAGACAGACAGATGATGAAAGAGTGCACCTCAAGAGCGTCAGTAC 998
OY 971 CTGCTGTGGGAAAGATCAAGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1030
Db 999 CTGCTGTGGGAAAGATCAAGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1058
OY 1031 GCCACCTCCAAAGCAGACAGGCTGCGGAAATGCGCTGCTGCTGCTGCTGCTGCTGCTG 1090
Db 1059 GCCACCTCCAAAGCAGACAGGCTGCGGAAATGCGCTGCTGCTGCTGCTGCTGCTGCTG 1118
OY 1091 AAGTTTCTGAGATCTCTCTGACAAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1150
Db 1119 AAGTTTCTGAGATCTCTCTGACAAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1178
OY 1151 AATGTGCTACCTTCAATTAAGCTGCTTCTTATAGGAAGGCTATGACACACTCTGGAG 1210
Db 1179 AATGTGCTACCTTCAATTAAGCTGCTTCTTATAGGAAGGCTATGACACACTCTGGAG 1238
OY 1211 AAAGCCAGGCTGAGCTGCGCAAGACCTTCCAGCAGCTCTGAGAGAGTCACTGGAGAG 1270
Db 1239 AAAGCCAGGCTGAGCTGCGCAAGACCTTCCAGCAGCTCTGAGAGAGTCACTGGAGAG 1298
OY 1271 CAGCTGAAGCCATGCTGAGAGTGGGCGCAAGTGGCTTCAAGCTCTGAGAGAGAGG 1330
Db 1299 CAGCTGAAGCCATGCTGAGAGTGGGCGCAAGTGGCTTCAAGCTCTGAGAGAGAGG 1358
OY 1331 CTCAAACCCAAAGAGACCAAGTGGTTAATTAAGTCAAGAGTGGCTTCAAGAGT 1390

Db 1359 CTCAAACCCAAAGAGACCAAGTGGTTAATTAAGTCAAGAGTGGCTTCAAGAGT 1418
OY 1391 AGGAACCTTAGAACCCAGAGACGCGAGAACAAAGTGCAGAGACGACACCAATGACTCT 1450
Db 1419 AGGAACCTTAGAACCCAGAGACGCGAGAACAAAGTGCAGAGACGACACCAATGACTCT 1478
OY 1451 GCTGCTTCTGAGTCCCGCCACCAAGCCCTCAAGACAAATAGCTATGGCGGAGAGAC 1510
Db 1479 GCTGCTTCTGAGTCCCGCCACCAAGCCCTCAAGACAAATAGCTATGGCGGAGAGAC 1538
OY 1511 CGAGGAGAGATGAGAGAGACCGAGAACGATGAGCTTGTGAAGTACCAACAAAGAGG 1570
Db 1539 CGAGGAGAGATGAGAGAGACCGAGAACGATGAGCTTGTGAAGTACCAACAAAGAGG 1598
OY 1571 AATCTGAAGACCGCTGTTGTCTGTGGAAGAAAGAACCTGTGTCTTCCACCCCTC 1630
Db 1599 AATCTGAAGACCGCTGTTGTCTGTGGAAGAAAGAACCTGTGTCTTCCACCCCTC 1658
OY 1631 TTGAGGGTGGGCTCTGTGACAGTTCGCGGATGCGCTCTAGAGCTCTTCTACATGAT 1690
Db 1659 TTGAGGGTGGGCTCTGTGACAGTTCGCGGATGCGCTCTAGAGCTCTTCTACATGAT 1718
OY 1691 GATGAGAGAGGCTATGAGTCTGACAGTTCGCGGATGCGCTCTAGAGCTCTTCTACAT 1750
Db 1719 GATGAGAGAGGCTATGAGTCTGACAGTTCGCGGATGCGCTCTAGAGCTCTTCTACAT 1778
OY 1751 TGCAGTAAACAAAGCTGCTGCAATGCTTGTGTGAGAGTGTGTGAGAGTGTGTGAG 1810
Db 1779 TGCAGTAAACAAAGCTGCTGCAATGCTTGTGTGAGAGTGTGTGAGAGTGTGTGAG 1838
OY 1811 GCAGGACAGCTGAGATGCCAAGCTGACAGAACCTGAGCTCTCTATATGCTGCTCT 1870
Db 1839 GCAGGACAGCTGAGATGCCAAGCTGACAGAACCTGAGCTCTCTATATGCTGCTCT 1898
OY 1871 CAGCGCTGACAGGCTGCTGCTCCGACGAGAAAGATGGAACATGCGCTGCAAGCTTC 1930
Db 1899 CAGCGCTGACAGGCTGCTGCTCCGACGAGAAAGATGGAACATGCGCTGCAAGCTTC 1958
OY 1931 TTCAGTACTGATCTGACCTGGAAGAAATTTGAGCCCAAGTGTGACCCAGCAATTCCT 1990
Db 1959 TTCAGTACTGATCTGACCTGGAAGAAATTTGAGCCCAAGTGTGACCCAGCAATTCCT 2018
OY 1991 GCAGCCAAAGAGGAGCCATAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2050
Db 2019 GCAGCCAAAGAGGAGCCATAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2078
OY 2051 TTGCTGCTCAAGGATTTGGATTTAAAGTGAAGTACATTCCTCGAAGTCTGTGCA 2110
Db 2079 TTGCTGCTCAAGGATTTGGATTTAAAGTGAAGTACATTCCTCGAAGTCTGTGCA 2138
OY 2111 GATCCATGCTGCTGAGGAACTGTAAGCATGAAAGGCAAGTCAATTAATGTAATGAGCTC 2170
Db 2139 GATCCATGCTGCTGAGGAACTGTAAGCATGAAAGGCAAGTCAATTAATGTAATGAGCTC 2198
OY 2171 CGGAAATATCCAAAGAAATATTTGAAGAGTGGGCGCTGCTGCTGCTGCTGCTGCTG 2230
Db 2199 CGGAAATATCCAAAGAAATATTTGAAGAGTGGGCGCTGCTGCTGCTGCTGCTGCTG 2258
OY 2231 AGCCCATGCAATGATCTCTTACGTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2290
Db 2259 AGCCCATGCAATGATCTCTTACGTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2318
OY 2291 GGAAGGCTTCTTCTGAGTTTACCACTTCTGTAATTAATCCCGCCCAAGAGAGGCGAC 2350
Db 2319 GGAAGGCTTCTTCTGAGTTTACCACTTCTGTAATTAATCCCGCCCAAGAGAGGCGAC 2378
OY 2351 AACCTGCAATCTTCTGAGTTTACCACTTCTGTAATTAATCCCGCCCAAGAGAGGCGAC 2410
Db 2379 AACCTGCAATCTTCTGAGTTTACCACTTCTGTAATTAATCCCGCCCAAGAGAGGCGAC 2438
OY 2411 GACATCTCAAGATTTCTGCTGATGTAAACCAAGTATGATGATGATGATGATGATGATG 2470
Db 2439 GACATCTCAAGATTTCTGCTGATGTAAACCAAGTATGATGATGATGATGATGATGATG 2498

QY 2471 GCTCACAGGGCCGGTACTTCTGGGGTAACCTACCCGGAATGAACAGCCCGTATGCT 2530
 |||||||
 Db 2499 GCTCACAGGGCCGGTACTTCTGGGGTAACCTACCCGGAATGAACAGCCCGTATGCT 2558
 |||||||
 QY 2531 TCAGAAATGATTAAGCTGAGCTGAGAGCTGCTGAGATTCACTAGAGACAGCAAGTTA 2590
 |||||||
 Db 2559 TCAGAAATGATTAAGCTGAGAGCTGAGAGCTGCTGAGATTCACTAGAGACAGCAAGTTA 2618
 |||||||
 QY 2591 AAGAAATGAGCAATTAACCAATGCAATCTCCATCAGACAGGGCCAAACACACTT 2650
 |||||||
 Db 2619 AAGAAATGAGCAATTAACCAATGCAATCTCCATCAGACAGGGCCAAACACACTT 2678
 |||||||
 QY 2651 TTCCCTAGTATCATGATGAGCAAGAGAGAGCTTTTGGTGCATCTGAGCTGAGCAAGATC 2710
 |||||||
 Db 2679 TTCCCTAGTATCATGATGAGCAAGAGAGAGCTTTTGGTGCATCTGAGCTGAGCAAGATC 2738
 |||||||
 QY 2711 TTCCGGCTTCCCTGCTACTACAGGAGCTGTCCAATGAGGCCGCGCCCTCAGAAAG 2770
 |||||||
 Db 2739 TTCCGGCTTCCCTGCTACTACAGGAGCTGTCCAATGAGGCCGCGCCCTCAGAAAG 2798
 |||||||
 QY 2771 CTGCTGGGAGAGTCCCTGAGTACCGGTATCAGAACCTGTTGGCCCTTGAAGGAC 2830
 |||||||
 Db 2799 CTGCTGGGAGAGTCCCTGAGTACCGGTATCAGAACCTGTTGGCCCTTGAAGGAC 2858
 |||||||
 QY 2831 TACTTTGCTGTAATAGTTCTACCCAGAGCTGGGAGCTCTCGTCAAGACCACTGACC 2890
 |||||||
 Db 2859 TACTTTGCTGTAATAGTTCTACCCAGAGCTGGGAGCTCTCGTCAAGACCACTGACC 2918
 |||||||
 QY 2891 AGAGTACCCCTCCCTGAAGGACACTGCTCCCTTTTAACTGACCTGCTGTGGGC 2950
 |||||||
 Db 2919 AGAGTACCCCTCCCTGAAGGACACTGCTCCCTTTTAACTGACCTGCTGTGGGC 2978
 |||||||
 QY 2951 CTGACATCATGTAACCTGAGTTCCTGCTGAGTGGAGCAAGCTCCTGGCCCTG 3010
 |||||||
 Db 2979 CTGACATCATGTAACCTGAGTTCCTGCTGAGTGGAGCAAGCTCCTGGCCCTG 3038
 |||||||
 QY 3011 CAGGGAGAGCCCGGTGCTCCCTCGTGTGACAGACTGAGACTGCTGCTTAAAGTACC 3070
 |||||||
 Db 3039 CAGGGAGAGCCCGGTGCTCCCTCGTGTGACAGACTGAGACTGCTGCTTAAAGTACC 3098
 |||||||
 QY 3071 CGGATGCTGCTCATGTTCTCTTACCTGAACCTTTAACTGAGTAACTAGTAAAGT 3130
 |||||||
 Db 3099 CGGATGCTGCTCATGTTCTCTTACCTGAACCTTTAACTGAGTAACTAGTAAAGT 3158
 |||||||
 QY 3131 GCGTCTCTTTAACCTGCTGAGTTATCATCAGAAAGTAAAGTAACTCAAAAAA 3190
 |||||||
 Db 3159 GCGTCTCTTTAACCTGCTGAGTTATCATCAGAAAGTAAAGTAACTCAAAAAA 3218
 |||||||
 QY 3191 CAACAACAAACAGAAACAAAAAACAACCTCAACAGCTCTCTTACTACTCAGT 3250
 |||||||
 Db 3219 CAACAACAAACAGAAACAAAAAACAACCTCAACAGCTCTCTTACTACTCAGT 3278
 |||||||
 QY 3251 TCATGCTGCAAAATCACTGAGATTTGTTTTAAGTAACCCGTGCTCCACATTTGCTGG 3310
 |||||||
 Db 3279 TCATGCTGCAAAATCACTGAGATTTGTTTTAAGTAACCCGTGCTCCACATTTGCTGG 3338
 |||||||
 QY 3311 AGATGCTATGTAATGTGGGCTGACATGAGCAAGTCAAGGGGCCAAAAAATTCC 3370
 |||||||
 Db 3339 AGATGCTATGTAATGTGGGCTGACATGAGCAAGTCAAGGGGCCAAAAAATTCC 3398
 |||||||
 QY 3371 CCTTCCGCCAGAGATATTTGAAGATGATTTTAACTGCTTCTGAGCACTTC 3430
 |||||||
 Db 3399 CCTTCCGCCAGAGATATTTGAAGATGATTTTAACTGCTTCTGAGCACTTC 3458
 |||||||
 QY 3431 CCTTCTTGTGTAACAAGGCTGAAGTCTGTTGCTTGTAGCAATTTCCAGAGATGATG 3490
 |||||||
 Db 3459 CCTTCTTGTGTAACAAGGCTGAAGTCTGTTGCTTGTAGCAATTTCCAGAGATGATG 3518
 |||||||
 QY 3491 ATGTCAGAGGAGATGACATCAACCTTTAGCGCTTTCCCTGAGAGGGCCCATGTGCG 3550
 |||||||
 Db 3519 ATGTCAGAGGAGATGACATCAACCTTTAGCGCTTTCCCTGAGAGGGCCCATGTGCG 3578
 |||||||

QY 3551 TAGTCTCACAGAGACTGAGATAGATTTTGGAGCTCAGAGAGCTGGTGCAGTGGCC 3610
 |||||||
 Db 3579 TAGTCTCACAGAGACTGAGATAGATTTTGGAGCTCAGAGAGCTGGTGCAGTGGCC 3638
 |||||||
 QY 3611 CTCTTCCAGGTGTAGGAGATACGAAGAGAGAGCTTAAAGAAATCATTCCTCCACTCC 3670
 |||||||
 Db 3639 CTCTTCCAGGTGTAGGAGATACGAAGAGAGAGCTTAAAGAAATCATTCCTCCACTCC 3698
 |||||||
 QY 3671 CTTCGCAAAATGAGGGGCCAGTCCCAACAGCTAGGTCCCAACACCCCTATTTCTC 3730
 |||||||
 Db 3699 CTTCGCAAAATGAGGGGCCAGTCCCAACAGCTAGGTCCCAACACCCCTATTTCTC 3758
 |||||||
 QY 3731 ATGAGAACTAGAGACCAAGACATGTTCCCTTATCTGAGCAGTGTGGGAACTA 3790
 |||||||
 Db 3759 ATGAGAACTAGAGACCAAGACATGTTCCCTTATCTGAGCAGTGTGGGAACTA 3818
 |||||||
 QY 3791 CAGTGAACACCTTGTGAGATGTTAAAGCTTTTACCACGATAGATGTGTTTTAA 3850
 |||||||
 Db 3819 CAGTGAACACCTTGTGAGATGTTAAAGCTTTTACCACGATAGATGTGTTTTAA 3878
 |||||||
 QY 3851 GGGGCTTTTATGAGGGGCATCACTGAGATTAAGAAAGTGCATTTCAAGAAATGCCATC 3910
 |||||||
 Db 3879 GGGGCTTTTATGAGGGGCATCACTGAGATTAAGAAAGTGCATTTCAAGAAATGCCATC 3938
 |||||||
 QY 3911 GTAATGTTTTTAAACACCTTTTACTAATTAACAGTGTCTATTTATAGAACAGACAC 3970
 |||||||
 Db 3939 GTAATGTTTTTAAACACCTTTTACTAATTAACAGTGTCTATTTATAGAACAGACAC 3998
 |||||||
 QY 3971 ACTTCTTTTATGACTCTCAGACTTCTATTTTCAATGTTACATTTTGTAACTGCA 4030
 |||||||
 Db 3999 ACTTCTTTTATGACTCTCAGACTTCTATTTTCAATGTTACATTTTGTAACTGCA 4058
 |||||||
 QY 4031 AGGTGGGCTTTTGTAACTCAGAGGTGGGAGAGACGCTGTTGTTCAAGATTG 4090
 |||||||
 Db 4059 AGGTGGGCTTTTGTAACTCAGAGGTGGGAGAGACGCTGTTGTTCAAGATTG 4118
 |||||||
 QY 4091 TCTCCACTGTTCTTAATTTTAAAGTGAAGATGACAGATGCCAGATTAACCTTCT 4150
 |||||||
 Db 4119 TCTCCACTGTTCTTAATTTTAAAGTGAAGATGACAGATGCCAGATTAACCTTCT 4178
 |||||||
 QY 4151 GGTGATTAAGTGTATTTCTTAATAAAAAAAAAAAAAAAAAAAAAA 4195
 |||||||
 Db 4179 GGTGATTAAGTGTATTTCTTAATAAAAAAAAAAAAAAAAAAAAAA 4223
 |||||||

RESULT 6
 AF151974
 LOCUS 4163 bp mRNA linear ROD 08-JUN-2000
 DEFINITION Mus musculus DNA cytosine-specific methyltransferase isoform 6
 (Dnmt3b) mRNA, complete cds.
 ACCESSION AF151974
 VERSION AF151974.1 GI:8347130
 KEYWORDS
 SOURCE Mus musculus.
 ORGANISM Mus musculus; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 4163)
 AUTHORS Yin, B., Chen, Y. T., Zhu, M., Luo, Y. J., Zhu, N., Xu, S. C., Wu, G. Y. and Shen, Y.
 TITLE Cloning of full-length Dnmt3b cDNA and its alternative splicing isoforms in mouse embryonic tissue
 JOURNAL 2 (bases 1 to 4163)
 REFERENCE Unpublished
 AUTHORS Yin, B., Chen, Y. T., Zhu, M., Luo, Y. J., Zhu, N., Xu, S. C., Wu, G. Y. and Shen, Y.
 TITLE Direct Submission
 JOURNAL Submitted (17-MAY-1999) Department of Biochemistry and Molecular Biology, Institution of Basic Medical Sciences, Chinese Academy of Medical Sciences, Dong Dan San Tiao 5, Beijing City 100005, P. R. China
 FEATURES
 source 1..4163
 location/Qualifiers

gene
CDS
/organism="Mus musculus"
/strain="KM"
/db_xref="taxon:10090"
/dev_stage="8-9 day old embryo"
1..4163
/gene="Dnmt3b"
297..2816
/gene="Dnmt3b"
/function="methylates cytosine in DNA"
/note="alternatively spliced"
/codon_start=1
/product="DNA cytosine-specific methyltransferase isoform 6"
/protein_1d="AA74520.1"
/db_xref="GI:8347131"
/translation="MKGDSRLHNEEGASCTEIIYNGFSDQSDTKDAPSYLVE
ALTEPCTPPTGRSSRLSKREVSLANTYQDMTGDDGDEVDGSDILMPK
LIRETKDRSESPAVRHSNGTSLERQASRLTRHOGHHVOEYVPEFATR
SRRRASASSTPMSPAVDMEVTPKSVSTPSVDLSQDQDEGMDITQVDAEILY
GSTEYDDKEFFGIGDLVWGKIKGFSWPMVAVSWKASKQAMPGRMVOFGDKF
SEISADKLVALGLFSOHFNLAFTNKLVSRYKAMHTLEKAVRAGTSSPESLED
OLKPLEWAGHGFPTGIEGLKPNKQENKSRRTTNDASASEPPKLTNSYSG
KDRGDEBSREHMASVEYTNKGNLEDRCLSCGKRPVFPFLFEGILQSCRRFLET
FMYDEDOYQSYCVCEGCEHLLCSNTSCRCFCVECELEVLVAGAGTAEADAKLOEFS
CYMCLPDRGHVLRKRDMMRLQDFETTPDLFEFEPKLYPAIPAKRRPRLVSL
FGIATGYLVKELGIVKERYIASEYCAESIAVGTGHEQIKVNDVRKLTKNIEE
WGPFLVGLSPCNDLSNVNPAKGLYESTGRLEFEEYHLLNTRPREGDNRPFEMF
ENVVAMKNDKDISREFLACNPYMDA.IKVSAAHARVFMGNI.LPQMRPYMAKNDKL
ELQDCLESRTAKIKVQITTKTSNIRGKNOLPPRYMNGKQDVLMTCELEIRGFP
AHYDYSMGSAKOKILGSMVSPYIRLRLPKDYACE"

BASE COUNT 1032 a 1074 c 1128 g 929 t
ORIGIN

Query Match 90.9% Score 3812.8; DB 10; Length 4163;
Best Local Similarity 98.3% Pzed. No. 0;
Matches 3878; Conservative 0; Mismatches 7; Indels 60; Gaps 1;
QY 251 GCCTCTTTCAAGAAACANTGAGAGGACAGCAGCATCTGATGAAAGAGGGTGCC 310
DB 279 GCCTCAGCAGAAACATGAGAGGAGAGCAGCAGCATCTGATGAAAGAGGGTGCC 338
QY 311 AGCGGATAGAGAGTGCATTATCGTTAATGGAACTTCAAGTCCCTCAGACAG 370
DB 339 AGCGGATAGAGAGTGCATTATCGTTAATGGAACTTCAAGTCCCTCAGACAG 398
QY 371 AAGGATGCTCCCTCAGCCAGCTCTTGAGAGCAATCTGCACAGAGCCAGTCTGCACACA 430
DB 399 AAGGATGCTCCCTCAGCCAGCTCTTGAGAGCAATCTGCACAGAGCCAGTCTGCACACA 458
QY 431 GAGACAGAGGCGGAGGTCAAGTCCCGGCTGTCTAAGAGGAGAGTCTCCAGCTTCTG 490
DB 459 GAGACAGAGGCGGAGGTCAAGTCCCGGCTGTCTAAGAGGAGAGTCTCCAGCTTCTG 518
QY 491 AATTACAGCAGAGCATGACAGAGATGAGAGATGAGATGAGATGAGATGAGATGAGAT 550
DB 519 AATTACAGCAGAGCATGACAGAGATGAGAGATGAGATGAGATGAGATGAGATGAGAT 578
QY 551 GGCTGATATTCTAATGCAAAAGCTCAGCCGTGAGACCAAGAGACCGAGCGCTCT 610
DB 579 GGCTGATATTCTAATGCAAAAGCTCAGCCGTGAGACCAAGAGACCGAGCGCTCT 638
QY 611 GAGAGCGGCGGTCTCCGAGCCGACATAGCAATGGGACCTCCAGCTTGAGAGGCAAGA 670
DB 639 GAAAGCGGCGGTCTCCGAGCCGACATAGCAATGGGACCTCCAGCTTGAGAGGCAAGA 698
QY 671 GCCTCCCCAGAAATCACCAGGTGCGGAGGGCCGACCAATGTGCGAGGATACCTGTG 730
DB 699 GCCTCCCCAGAAATCACCAGGTGCGGAGGGCCGACCAATGTGCGAGGATACCTGTG 758
QY 731 GAGTTCCGCGTACAGAGTCTCGAGAGAGTCTGAGAGAGTCTGAGAGAGAGAGAGAG 790
DB 759 GAGTTCCGCGTACAGAGTCTCGAGAGAGTCTGAGAGAGTCTGAGAGAGAGAGAGAG 818

QY 791 TCATCCCTGCGACAGTCTCATGAAAGAGTACACCTAAGAGCGTCACTACCCCA 850
DB 819 TCATCCCTGCGACAGTCTCATGAAAGAGTACACCTAAGAGCGTCACTACCCCA 878
QY 851 TCAGTTGACTTGAGCCAGAGATGAGATGAGAGGATGAGATGAGATGAGATGAGATGAGAT 910
DB 879 TCAGTTGACTTGAGCCAGAGATGAGATGAGAGGATGAGATGAGATGAGATGAGATGAGAT 938
QY 911 GAGAGCAAGATGAGAGACAGACAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 970
DB 939 GAGAGCAAGATGAGAGACAGACAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 998
QY 971 CTGCTGTGGGAAAGATCAAGGCTTCTCTGTGTGGCTCCCATGCTGTGTCTGGAAA 1030
DB 999 CTGCTGTGGGAAAGATCAAGGCTTCTCTGTGTGGCTCCCATGCTGTGTCTGGAAA 1058
QY 1031 GCCACCTCAAGCGACAGGCGCATGCCGGAATGCCGTGGTACAGTGGTTGGTGTATGCC 1090
DB 1059 GCCACCTCAAGCGACAGGCGCATGCCGGAATGCCGTGGTACAGTGGTTGGTGTATGCC 1118
QY 1091 AAGTTCTGAGATCTGCTGCAAAACTGTGGCTGTGGGCTGTTCACGACACTTT 1150
DB 1119 AAGTTCTGAGATCTGCTGCAAAACTGTGGCTGTGGGCTGTTCACGACACTTT 1178
QY 1151 AATCTGGCTACCTCATTAAGTGTGTTCTTATAGAAAGGCGATGACACACTGTGGAG 1210
DB 1179 AATCTGGCTACCTCATTAAGTGTGTTCTTATAGAAAGGCGATGACACACTGTGGAG 1238
QY 1211 AAGCGAGGTGAGGCTGCGCAAGACTCTGACAGACTCTGAGAGAGTCACTGAGAGAC 1270
DB 1239 AAGCGAGGTGAGGCTGCGCAAGACTCTGACAGACTCTGAGAGAGTCACTGAGAGAC 1298
QY 1271 CAGCTGAAGCCCATGCTGAGTGGGCGCCAGGCTGCTCAAGCTACTGGGATTCAGAGGC 1330
DB 1299 CAGCTGAAGCCCATGCTGAGTGGGCGCCAGGCTGCTCAAGCTACTGGGATTCAGAGGC 1358
QY 1331 CTCAAAACCAAGAAAGAACCAACGATGTTAATAGTCAAGAGTGGCTGTTCAAGACT 1390
DB 1359 CTCAAAACCAAGAAAGAACCAACGATGTTAATAGTCAAGAGTGGCTGTTCAAGACT 1381
QY 1391 AAGAACTTGAACCAAGAGACGCGAGAAACAAAGTCAAGAGCGACACCAATGACTCT 1450
DB 1382 -----CGAGAAACAAAGTCAAGAGCGACACCAATGACTCT 1418
QY 1451 GCTGCTTGTAGTCCCCCCCCACCAAGCGCTCAAGACAAATAGCTATGGCGGAAAGAC 1510
DB 1419 GCTGCTTGTAGTCCCCCCCCACCAAGCGCTCAAGACAAATAGCTATGGCGGAAAGAC 1478
QY 1511 CGAGGAGAGATGAGAGACCGAGAACGATGGCTTGTGAAGTCAACCAACCAAGAGGC 1570
DB 1479 CGAGGAGAGATGAGAGACCGAGAACGATGGCTTGTGAAGTCAACCAACCAAGAGGC 1538
QY 1571 AATCTGGAAGCGCGCTTGTGCTGGAAGAAAGAAACCTGTGCTTCCACCCCTC 1630
DB 1539 AATCTGGAAGCGCGCTTGTGCTGGAAGAAAGAAACCTGTGCTTCCACCCCTC 1598
QY 1631 TTTGAGGCTGGGCTCTGTCAAGATTGCGGAGTGCCTTCTAAGACTCTTCAATGAT 1690
DB 1599 TTTGAGGCTGGGCTCTGTCAAGATTGCGGAGTGCCTTCTAAGACTCTTCAATGAT 1658
QY 1691 GATGAGAGCGGCTATCGTCTACTGACACCGTGTGCTGAGAGGCGCGTGAATGCTGTG 1750
DB 1659 GATGAGAGCGGCTATCGTCTACTGACACCGTGTGCTGAGAGGCGCGTGAATGCTGTG 1718
QY 1751 TGCATTAACACAACTCTGTCAGATGCTTCTGTGAGAGTCTGAGAGGCTGTGGGAGC 1810
DB 1719 TGCATTAACACAACTCTGTCAGATGCTTCTGTGAGAGTCTGAGAGGCTGTGGGAGC 1778
QY 1811 GCAGGCAAGCTGAGATGCAAGCTGCGAGAAACCTGGAACCTGATATGTCCTCCT 1870
DB 1779 GCAGGCAAGCTGAGATGCAAGCTGCGAGAAACCTGGAACCTGATATGTCCTCCT 1838
QY 1871 CAGCGCTGCCATGGGCTCTCCGACGAGGAAAGATTGMAACATGGCGCTCGAAGACTTC 1930

QY 4091 TCTCAGCTGTTTCTTAATTTTGTAGTCGAAGATGACAGATGCCAGATTACTTCT 4150
DB 4059 TCTCAGCTGTTTCTTAATTTTGTAGTCGAAGATGACAGATGCCAGATTACTTCT 4118
QY 4151 GGTGATTAAGTGTATTTCTTAATAAAAAAAAAAAAAAAAAAAAAA 4195
DB 4119 GGTGATTAAGTGTATTTCTTAATAAAAAAAAAAAAAAAAAAAAAA 4163

RESULT 7
AY078427 4006 bp mRNA linear ROD 01-JUL-2002
LOCUS Mus musculus DNA cytosine methyltransferase 3b6 (Dnmt3b6) mRNA,
DEFINITION complete cds; alternatively spliced.
ACCESSION AY078427
VERSION AY078427.1 GI:21655120
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
UTHERS Chen, T., Ueda, Y. and Li, E.
TITLE Dnmt3a2 encoded by transcripts from an intronic promoter of Dnmt3a
is expressed abundantly in ES cells and germ cells
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4006)
AUTHORS Chen, T., Ueda, Y. and Li, E.
TITLE Direct Submision
JOURNAL Submitted (08-FEB-2002) Cardiovascular Research Center,
Massachusetts General Hospital, 149 13th Street, Charlestown, MA
02129, USA

FEATURES
source
1..4006 Location/Qualifiers
/organism="Mus musculus"
/strain="129Sv/evj"
/db_xref="taxon:10090"
/chromosome="2"
/map="84.0 cM"
1..4006
/gene="Dnmt3b6"
269..2659
/gene="Dnmt3b6"
/function="de novo methylation"
/note="contains PHD and PWWP domains; alternatively
spliced"
/codon_start=1
/product="DNA cytosine methyltransferase 3b6"
/protein_id="AAL85481.1"
/db_xref="GI:21655121"
translation="MKGDSRLNLEEGASGYEECIYNGNPSOSSDTKADAPSYLE
LICEPYCTPETGRSSRLSKREYSILNTODMTGDGDDRDNDVDDGSDILMPK
LIREKIDRTRESPAVTRRHNGTSILRORASPTLRDGRDNDVDDGSDILMPK
SRRRASSASTWSSPAVDPEMTKPSYTSVLSDDGQDGMOTDTPDASBR
GDSREYDDKPEGLDLYGCKIKGSMWPMWPMWPMWPMWPMWPMWPMWPMW
SEIADLVAGLPSQHPNLATPNKLVSRKAMTLEKAVRAGKTFSSPESGDR
OLKPLPRACCKRPTGIDGLKPKKQRYNKSRYKSDSNLEPRRERKRNFTN
DSASESPFRKRLTNSYGGDKGDEDESRMASVYNNKNGLEDCLSCKRNFTN
FHPLEGLDSCSDRFLFTFTYDDEGQSTVCCBGRLLCSNTSCRCFEC
LEVLAGTADKLOEPWSCYCLPQRCHGLRKRDMNRLQDEFTFDPLEFEP
FKLPAIPAKRRPRLVSLFDGIALTVLKEGLIYEVYIASEVCASIAVGYKH
EGQIKYVNDVRKITKNIEMGPFDLVIGGSPCNDLSNVNPARGKIEGGRLEF
HLNLTNRKGNRPFFMFEENVAMKVNDRKDISRLACNPVMDIAIVYSAHARY
FMGNLPGMNNIFGEPAYHYDVSNGGAKRKLGRSWSVPIRHLFAPLKYFACE"

BASE COUNT 1001 a 1017 c 1075 g 913 t
ORIGIN

Query Match 90.8%; Score 3807; DB 10; Length 4006;
Best Local Similarity 95.5%; Pred. No. 0;
Matches 4006; Conservative 0; Mismatches 0; Indels 189; Gaps 1;

QY 1 GAATTCGGGGCGGGGTTAAGCGGCCCAAGTAACGTACGACGAGATCGGCGCGGA 60

DB 1 GAATTCGGGGCGGGGTTAAGCGGCCCAAGTAACGTACGACGAGATCGGCGCGGA 60
QY 61 GAATTCGGGGCGGGGTTAAGCGGCCCAAGTAACGTACGACGAGATCGGCGCGGA 120
DB 61 GAATTCGGGGCGGGGTTAAGCGGCCCAAGTAACGTACGACGAGATCGGCGCGGA 120
QY 121 GCGGGCTACAGACGCTTCCAGACAGAGCCGCTGAGGCTGTGTCAGACCTTGAACCC 180
DB 121 GCGGGCTACAGACGCTTCCAGACAGAGCCGCTGAGGCTGTGTCAGACCTTGAACCC 180
QY 181 TGAGTATATACCTTTTCAGACAGCGGGGATCTCCCTCCCATCATAGTCCCTGGGAC 240
DB 181 TGAGTATATACCTTTTCAGACAGCGGGGATCTCCCTCCCATCATAGTCCCTGGGAC 240
QY 241 CAAATTCAGAGGCTTTCTTTCAGGAAACAAATGAAGGAGACAGACATCTGAATAGA 300
DB 241 CAAATTCAGAGGCTTTCTTTCAGGAAACAAATGAAGGAGACAGACATCTGAATAGA 300
QY 301 AGAGGCTCCAGCGGATGAGAGTGCATATCGTTAATGGAACTTCAGTACAGTC 360
DB 301 AGAGGCTCCAGCGGATGAGAGTGCATATCGTTAATGGAACTTCAGTACAGTC 360
QY 361 CTCACACAGAAAGATGCTCCCTACACCCCATCTTGGAGGCATCTGCACAGAGCACT 420
DB 361 CTCACACAGAAAGATGCTCCCTACACCCCATCTTGGAGGCATCTGCACAGAGCACT 420
QY 421 CTGCACACAGAGACAGAGCGGCGGCTGACAGCTCCGCTGTCTAAGAGGAGTCTC 480
DB 421 CTGCACACAGAGACAGAGCGGCGGCTGACAGCTCCGCTGTCTAAGAGGAGTCTC 480
QY 481 CAGCCTTGTGAATTAACAGCAGACATGACAGAGATGAGACAGATATATATAGTAA 540
DB 481 CAGCCTTGTGAATTAACAGCAGACATGACAGAGATGAGACAGATATATATAGTAA 540
QY 541 TGATGGAGTGGCTGTGATATTTCTAATGCAAAAGCTCACCCGTGAGACCAAGCACCG 600
DB 541 TGATGGAGTGGCTGTGATATTTCTAATGCAAAAGCTCACCCGTGAGACCAAGCACCG 600
QY 601 GAGCGGCTGTGAAAGCCCGGCTGCGGACCGGACATGCAATGGAGCTCCAGCTTGA 660
DB 601 GAGCGGCTGTGAAAGCCCGGCTGCGGACCGGACATGCAATGGAGCTCCAGCTTGA 660
QY 661 GAGCGGCTGTGAAAGCCCGGACCGGACATGCAATGGAGCTCCAGCTTGA 720
DB 661 GAGCGGCTGTGAAAGCCCGGACCGGACATGCAATGGAGCTCCAGCTTGA 720
QY 721 GTACCTGTGAGTTTCCGGCTACAGTCTGTGAGAGCTGAGACATCGTTCACCAAG 780
DB 721 GTACCTGTGAGTTTCCGGCTACAGTCTGTGAGAGCTGAGACATCGTTCACCAAG 780
QY 781 CAGCGCATGATGATCCCGTCCAGCGCTGACATGCAATGCAATGCAATGCAATGCAAT 840
DB 781 CAGCGCATGATGATCCCGTCCAGCGCTGACATGCAATGCAATGCAATGCAATGCAAT 840
QY 841 CAGTACCCCATCACTTACTTACCTTACAGCAGATGAGATGAGAGGATATGATACACACA 900
DB 841 CAGTACCCCATCACTTACTTACCTTACAGCAGATGAGATGAGAGGATATGATACACACA 900
QY 901 GGTGATGACAGAGAGCAGATGAGACAGCAGACAGATATCAGATATTAAGATTTGG 960
DB 901 GGTGATGACAGAGAGCAGATGAGACAGCAGACAGATATCAGATATTAAGATTTGG 960
QY 961 AATAGTGTACCTGTGTGGGGAAGATCAAGGGCTTCTCTGTGTGCTCCATGTGTGT 1020
DB 961 AATAGTGTACCTGTGTGGGGAAGATCAAGGGCTTCTCTGTGTGCTCCATGTGTGT 1020
QY 1021 GTCTGTGAAAGCAACCTCCCAAGGAGACATGCGGGAATGCGGTACAGTGGTT 1080
DB 1021 GTCTGTGAAAGCAACCTCCCAAGGAGACATGCGGGAATGCGGTACAGTGGTT 1080
QY 1081 TGCTGATGCGCAATTTTCTGAGATCTCTGACACAAATGCTGCTGTGGGCTGTTCAG 1140
DB 1081 TGCTGATGCGCAATTTTCTGAGATCTCTGACACAAATGCTGCTGTGGGCTGTTCAG 1140

Db 1081 TGGTATGGCAGTTTTCTGATCTGCTGACMAACTGGTGGCTCTGGGGCTGTTCAG 1140
Qy 1141 CCAGACATTTAATGCTACCTCAATAGCTGGTTCTTATGSAAGGCCATGTCCA 1200
Db 1141 CCAGACATTTAATGCTACCTCAATAGCTGGTTCTTATGSAAGGCCATGTCCA 1200
Qy 1201 CACTCTGAGAAACCCAGGGGTGAGCTGGCAAGACCTTCTCCAGAGTCTCTGAGATC 1260
Db 1201 CACTCTGAGAAACCCAGGGGTGAGCTGGCAAGACCTTCTCCAGAGTCTCTGAGATC 1260
Qy 1261 ACTGAGAGCAGCTGGAAGCCCATGCTGAGTGGGCCAGCGTGTCAAGCTTACTGG 1320
Db 1261 ACTGAGAGCAGCTGGAAGCCCATGCTGAGTGGGCCAGCGTGTCAAGCTTACTGG 1320
Qy 1321 GATGAGGGGCTCAAAACCAAGAAAGCAACAGTGTATATAGTGAAGGTGGCTGG 1380
Db 1321 GATGAGGGGCTCAAAACCAAGAAAGCAACAGTGTATATAGTGAAGGTGGCTGG 1380
Qy 1381 TTCAGACAGTAGAACTTAGAACCCAGAGACGGAGAACMAAAGTGAAGACGACAGAC 1440
Db 1381 TTCAGACAGTAGAACTTAGAACCCAGAGACGGAGAACMAAAGTGAAGACGACAGAC 1440
Qy 1441 CAATGACTCTGCTCTTCTGAGTCCGCCCAAGCGCTCAAGACAAATAGCTATGG 1500
Db 1441 CAATGACTCTGCTCTTCTGAGTCCGCCCAAGCGCTCAAGACAAATAGCTATGG 1500
Qy 1501 CGGGAAGAGCCGAGGGAGGTAGAGAGCCGAAACGATGGCTTCTGAGTACACAA 1560
Db 1501 CGGGAAGAGCCGAGGGAGGTAGAGAGCCGAAACGATGGCTTCTGAGTACACAA 1560
Qy 1561 CAACAGAGGCAATCTGGAAGACCGCTTTTCTCTGGAAGAAGAACCTGTGCTCT 1620
Db 1561 CAACAGAGGCAATCTGGAAGACCGCTTTTCTCTGGAAGAAGAACCTGTGCTCT 1620
Qy 1621 CCACCCCTCTTTAGGGGTGGGCTCTGCAAGATTTGGCGGATGCTCTTCTAGAGCTCT 1680
Db 1621 CCACCCCTCTTTAGGGGTGGGCTCTGCAAGATTTGGCGGATGCTCTTCTAGAGCTCT 1680
Qy 1681 CTACATGATGATGAGAGCGGCTATCAGTCTCTACTGACCGTGGCTGTGAGGCGCTGA 1740
Db 1681 CTACATGATGATGAGAGCGGCTATCAGTCTCTACTGACCGTGGCTGTGAGGCGCTGA 1740
Qy 1741 ACTGCTCTGTGAGTAACACACAGCTGCTGAGATGCTGTGTGAGTCTCTGAGAGT 1800
Db 1741 ACTGCTCTGTGAGTAACACACAGCTGCTGAGATGCTGTGTGAGTCTCTGAGAGT 1800
Qy 1801 GCTGTGGGGCAGGACAGCTGAGATGCCAAGCTGACGAGAACCTGTGAGCTCTATAT 1860
Db 1801 GCTGTGGGGCAGGACAGCTGAGATGCCAAGCTGACGAGAACCTGTGAGCTCTATAT 1860
Qy 1861 GTGCTCCCTCAGGCGGCTAGGGGTCTCCGAGCGAGGAAAGTGGAACTAGCGCT 1920
Db 1861 GTGCTCCCTCAGGCGGCTAGGGGTCTCCGAGCGAGGAAAGTGGAACTAGCGCT 1920
Qy 1921 GCAAGACTTCTTCACTCTGATGCTGAGTGAAGAAATTTGAGCACCAGTGTGACCC 1980
Db 1921 GCAAGACTTCTTCACTCTGATGCTGAGTGAAGAAATTTGAGCACCAGTGTGACCC 1980
Qy 1981 AGCAATTCCTGACGCAAAAGAGGCCCATTAGAGTCTCTCTCTGTTGATGGAATGTC 2040
Db 1981 AGCAATTCCTGACGCAAAAGAGGCCCATTAGAGTCTCTCTCTGTTGATGGAATGTC 2040
Qy 2041 AAGGGGACTTGTGCTCAAGSAGTTGGGTATTAAGTGGAAAGTCAATGGCTCGCA 2100
Db 2041 AAGGGGACTTGTGCTCAAGSAGTTGGGTATTAAGTGGAAAGTCAATGGCTCGCA 2100
Qy 2101 AGTCTGTGACAGTCAATGCTGTGGAACTGTTAAGCATGAAGGCCAGATCAATATGT 2160
Db 2101 AGTCTGTGACAGTCAATGCTGTGGAACTGTTAAGCATGAAGGCCAGATCAATATGT 2160
Qy 2161 CAATGAGCTCGGAAATACCAAGAAATATTTGAAGAGTGGGCCCTGTGACTGGT 2220
Db 2161 CAATGAGCTCGGAAATACCAAGAAATATTTGAAGAGTGGGCCCTGTGACTGGT 2220

Qy 2221 GATTGGTGAAGCCCATGCAATGATCTCTTAACGTCAATCTGCGCCAAAGGTTATA 2280
Db 2221 GATTGGTGAAGCCCATGCAATGATCTCTTAACGTCAATCTGCGCCAAAGGTTATA 2280
Qy 2281 TGAAGGCAACAGAGGCTCTTCTTCTGAGTTTACCAGTGTGCTGAATATACCCGCCAA 2340
Db 2281 TGAAGGCAACAGAGGCTCTTCTTCTGAGTTTACCAGTGTGCTGAATATACCCGCCAA 2340
Qy 2341 GGAGGGGCAACAGCTGCTCTTCTGAGTGTGGAATGTGTGGCCATGAAGTGA 2400
Db 2341 GGAGGGGCAACAGCTGCTCTTCTGAGTGTGGAATGTGTGGCCATGAAGTGA 2400
Qy 2401 TGACAGAAAGACATCTCAAGATCTCGGCACTGTAACCAAGTATGATGCTGATCAA 2460
Db 2401 TGACAGAAAGACATCTCAAGATCTCGGCACTGTAACCAAGTATGATGCTGATCAA 2460
Qy 2461 GGTGTCTGCTGCTACAGAGGCGCGTACTTCTGGGGTAACTTACCAGGAATGAACAGGCC 2520
Db 2461 GGTGTCTGCTGCTACAGAGGCGCGTACTTCTGGGGTAACTTACCAGGAATGAACAGGCC 2520
Qy 2521 CGTATGGCTTCAAAAGATGATTAAGCTGAGCTCAGSACTGCTGGAGTTCACTAGGAC 2580
Db 2521 CGTATGGCTTCAAAAGATGATTAAGCTGAGCTCAGSACTGCTGGAGTTCACTAGGAC 2580
Qy 2581 AGCAAGTTAAGAAAGTGCAGACAAATTAACCAAGTGCMAACTCATCAGACAGAGGCA 2640
Db 2581 AGCAAGTTAAGAAAGTGCAGACAAATTAACCAAGTGCMAACTCATCAGACAGAGGCA 2640
Qy 2641 AAMCAGCTTTCCCTGATGATGAATGGCAAGAGACGTTTTGTGTGCTGACTGAGCT 2700
Db 2641 AAMCAGCTTTCCCTGATGATGAATGGCAAGAGACGTTTTGTGTGCTGACTGAGCT 2700
Qy 2701 CGAAAGATCTTGGCTTCCCTGCTACCTACACGAGAGTGTCCAAACATGGCCCGGCGC 2760
Db 2701 CGAAAGATCTTGGCTTCCCTGCTACCTACACGAGAGTGTCCAAACATGGCCCGGCGC 2760
Qy 2761 ---AGATCTTCCGCTTCCCTGCTACCTACACGAGAGTGTCCAAACATGGCCCGGCGC 2820
Db 2761 ---AGATCTTCCGCTTCCCTGCTACCTACACGAGAGTGTCCAAACATGGCCCGGCGC 2820
Qy 2821 CCGTACAGAGCTGTGGGAGAGTCTGGAAGTACCGGCTCATGACAGACCTGTTGGCCC 2880
Db 2821 CCGTACAGAGCTGTGGGAGAGTCTGGAAGTACCGGCTCATGACAGACCTGTTGGCCC 2880
Qy 2881 CTTGAAGACTTACTTGGCTGTGAATAGTTTACCCAGAGCTGGGAGCTCTCGGTGCA 2940
Db 2881 CTTGAAGACTTACTTGGCTGTGAATAGTTTACCCAGAGCTGGGAGCTCTCGGTGCA 2940
Qy 2941 GCGAGTGGCCAGAGTCAACCCCTCCCTGAAGGACACCTGCTGCCCTTTTACCTCAC 2991
Db 2941 GCGAGTGGCCAGAGTCAACCCCTCCCTGAAGGACACCTGCTGCCCTTTTACCTCAC 2991
Qy 2991 TGTGTGGGGCTCAATCACTGTAACCTGAGCTTCTCTGCTCACTGAGTGGAGACAGAGCTC 3051
Db 2991 TGTGTGGGGCTCAATCACTGTAACCTGAGCTTCTCTGCTCACTGAGTGGAGACAGAGCTC 3051
Qy 3051 CTGGCCCTTGAAGGAGAGCCCGGAGTCCCTCGGTGAGCAGAGCTGAGACTGGTGGCT 3111
Db 3051 CTGGCCCTTGAAGGAGAGCCCGGAGTCCCTCGGTGAGCAGAGCTGAGACTGGTGGCT 3111
Qy 3111 TAGAGTACCGCGGATGCTGATGCTTCTTACCTGAAACCTTAAACCTTGAAGTGA 3171
Db 3111 TAGAGTACCGCGGATGCTGATGCTTCTTACCTGAAACCTTAAACCTTGAAGTGA 3171
Qy 3171 TAGAGTACCGCGGATGCTGATGCTTCTTACCTGAAACCTTAAACCTTGAAGTGA 3231
Db 3171 TAGAGTACCGCGGATGCTGATGCTTCTTACCTGAAACCTTAAACCTTGAAGTGA 3231
Qy 3231 GTAGTACCGCGGATGCTGATGCTTCTTACCTGAAACCTTAAACCTTGAAGTGA 3291
Db 3231 GTAGTACCGCGGATGCTGATGCTTCTTACCTGAAACCTTAAACCTTGAAGTGA 3291
Qy 3291 GTAGTACCGCGGATGCTGATGCTTCTTACCTGAAACCTTAAACCTTGAAGTGA 3351
Db 3291 GTAGTACCGCGGATGCTGATGCTTCTTACCTGAAACCTTAAACCTTGAAGTGA 3351
Qy 3351 GTAGTACCGCGGATGCTGATGCTTCTTACCTGAAACCTTAAACCTTGAAGTGA 3411
Db 3351 GTAGTACCGCGGATGCTGATGCTTCTTACCTGAAACCTTAAACCTTGAAGTGA 3411

```

OY 3301 CATTTCGTGAGATGCTATTTGTAATGTGGCTCAGATGACAGCTCAAGGGGCCAA 3360
DB 3112 CATTTCGTGAGATGCTATTTGTAATGTGGCTCAGATGACAGCTCAAGGGGCCAA 3171
OY 3361 AAAAATTCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 3420
DB 3172 AAAAATTCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 3231
OY 3421 TGGCAGCTTCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 3480
DB 3232 TGGCAGCTTCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 3291
OY 3481 CAGATGATGATGTAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 3540
DB 3292 CAGATGATGATGTAGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 3351
OY 3541 CCCATGCTGCTAGTCTCAGAGATGATGATGATGATGATGATGATGATGATGATGAT 3600
DB 3352 CCCATGCTGCTAGTCTCAGAGATGATGATGATGATGATGATGATGATGATGATGAT 3411
OY 3601 TGGAGTGGCCCTCTTCCAGGTGTGAGAGATGATGATGATGATGATGATGATGAT 3660
DB 3412 TGGAGTGGCCCTCTTCCAGGTGTGAGAGATGATGATGATGATGATGATGATGAT 3471
OY 3661 CCCACTCCCTCTTCCAGGTGTGAGAGATGATGATGATGATGATGATGATGATGATGAT 3720
DB 3472 CCCACTCCCTCTTCCAGGTGTGAGAGATGATGATGATGATGATGATGATGATGATGAT 3531
OY 3721 CTAGTTCCTCATGAGAACTAGACAGACAGACAGACAGACAGACAGACAGACAGAC 3780
DB 3532 CTAGTTCCTCATGAGAACTAGACAGACAGACAGACAGACAGACAGACAGACAGAC 3591
OY 3781 TGGGGAACATGAGTAAACCTCTGAGATGATGATGATGATGATGATGATGATGATGAT 3840
DB 3592 TGGGGAACATGAGTAAACCTCTGAGATGATGATGATGATGATGATGATGATGATGAT 3651
OY 3841 GTGTTTTTAAAGGGGCTTTTAAAGGGGCTTTTAAAGGGGCTTTTAAAGGGGCTTTT 3900
DB 3652 GTGTTTTTAAAGGGGCTTTTAAAGGGGCTTTTAAAGGGGCTTTTAAAGGGGCTTTT 3711
OY 3901 AATGTCATGCTAATGCTTTTAAACCTTTTAAACCTTTTAAACCTTTTAAACCTTTT 3960
DB 3712 AATGTCATGCTAATGCTTTTAAACCTTTTAAACCTTTTAAACCTTTTAAACCTTTT 3771
OY 3961 AGCAGCAACACTTTTAAAGCTGCTGAGCTGATGATGATGATGATGATGATGATGAT 4020
DB 3772 AGCAGCAACACTTTTAAAGCTGCTGAGCTGATGATGATGATGATGATGATGATGAT 3831
OY 4021 GTAACTCGCAAGCTGAGGCTTTTAACTTCAAGCTGAGGCTGAGGCTGAGGCTGAG 4080
DB 3832 GTAACTCGCAAGCTGAGGCTTTTAACTTCAAGCTGAGGCTGAGGCTGAGGCTGAG 3891
OY 4081 CAACATGCTGCTCAGCTGCTTTTAACTTCAAGCTGAGGCTGAGGCTGAGGCTGAG 4140
DB 3892 CAACATGCTGCTCAGCTGCTTTTAACTTCAAGCTGAGGCTGAGGCTGAGGCTGAG 3951
OY 4141 TTACCTTTGCTGATTAAGTTGATTTCTCTAATAAAAAAAAAAAAAAAAAAAAA 4195
DB 3952 TTACCTTTGCTGATTAAGTTGATTTCTCTAATAAAAAAAAAAAAAAAAAAAAA 4006

```

```

REFERENCE 1 (bases 1 to 4149)
AUTHORS Yin,B., Chen,X.T., Zhu,M., Luo,Y.J., Zhu,N., Xu,S.C., Wu,G.Y. and
          Shen,Y.
TITLE Cloning of full-length Dnmt3b cDNA and its alternative splicing
        isoforms in mouse embryonic tissue
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4149)
AUTHORS Yin,B., Chen,X.T., Zhu,M., Luo,Y.J., Zhu,N., Xu,S.C., Wu,G.Y. and
          Shen,Y.
TITLE Direct Submission
JOURNAL Submitted (17-MAY-1999) Department of Biochemistry and Molecular
          Biology, Institution of Basic Medical Sciences, Chinese Academy of
          Medical Sciences, Dong Dan San Tiao 5, Beijing City 100005,
          P.R.China
FEATURES
Source 1. 4149
        /organism="Mus musculus"
        /strain="KM"
        /db_xref="taxon:10090"
        /dev_stage="8-9 day old embryo"
        1. 4149
        /gene="Dnmt3b"
        /gene="Dnmt3b"
        /function="methylates cytosine in DNA"
        /note="alternatively spliced"
        /codon_start=1
        /product="DNA cytosine-specific methyltransferase isoform
        4"
        /protein_id="AAE74518.1"
        /db_xref="gi:8347126"
        /translation="MKGDSRLHNEEGASGYEECIIVNGNFSPOSSDITKADSPVLE
        AICREPCPEPTGRRSSRLSKREBSLAWTODMKGDSDRDDEVDGNSDILMPK
        LTRREKPTRESPEVARTRHNSGTSLEKORASPRTRRORGRHHVOYQEPAPATR
        SRBRASASAPSWSSPASVDPMEVETPKSYSTSVLSODGDRGDMOTVOYDASLY
        GDSREYDDKPERGIDLVNGKIKGFSMPMAVYKATSRQAMPCMPHVOYFGCKP
        SEIADLVILGFSPHPLATPKNLKYSRAMVHTLEKARYACGTSSSCSELED
        QKRLDLMAHGGKPTGIDGLKPKKQDPYKMSKVRSDSRNLEPRRERAKRRRTTN
        DNASSEPTPRKATNSYSGKDRGDEBSERMASEYTNKNGLEDRCLSCGRKPVY
        FHLFEGGLQSCDRFLFTFYDEDTGISTVCCGSEHLICSNISCCGFVEEC
        LEVLVAGTAEKDLDEPMSCTKLPORCHVEYLRKKDMRMLQDFTFDPLEVEBP
        KPLPALPAKRRPRIVLSLEFGIATGLVLEKIVERTIASEVCASIVAGVKA
        EGQIRYVNDKRTKRNKNIEMWGPEDLVIGSSPCDLSNVPARKGLIYETGVLFEFY
        HILNVTYPRKEDNRPFPMPFENVAMKVNDRKDISRFLACNPVMDAIVSAHARY
        FGMNLPQMNRIFFEPAHYTDVSNMGRAROKLLGRMSVPIVRIHLFAPLDYFACE"
BASE COUNT 1013 a 1082 c 1123 g 931 t
ORIGIN
Query Match 90.5%; Score 3794.8; DB 10; Length 4149;
Best Local Similarity 95.4%; Pred. No. 0;
Matches 3995; Conservative 0; Mismatches 2; Indels 189; Gaps 1;

```

```

OY 10 GCGCGGGGTTAAACGCGCCCAATAAAGCTAGCGAGGATGCGCGCGGAGATTGCGCA 69
DB 153 GCGCGGGGTTAAACGCGCCCAATAAAGCTAGCGAGGATGCGCGCGGAGATTGCGCA 212
OY 70 ACCGCAACCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 129
DB 213 ACCGCAACCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 272
OY 130 CAGCCAGCTCAGACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 189
DB 273 CAGCCAGCTCAGACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 332
OY 190 TACCTTTCCAGACGCGGATCTCCCTCCCATCATAGTGCCTTGGACCAATTCAG 249
DB 333 TACCTTTCCAGACGCGGATCTCCCTCCCATCATAGTGCCTTGGACCAATTCAG 392
OY 250 GGCCTTTCTTCCAGAAACATGAAGGAGACAGACAGATCTGAATGAAGAGAGGTGC 309
DB 393 GGCCTTTCTTCCAGAAACATGAAGGAGACAGACAGATCTGAATGAAGAGAGGTGC 452
OY 310 CAGCGGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 369

```

```

RESULT 8
AF151972 4149 bp mRNA linear ROD 08-JUN-2000
LOCUS AF151972 Mus musculus DNA cytosine-specific methyltransferase isoform 4
DEFINITION (Dnmt3b) mRNA, complete cds.
ACCESSION AF151972.1 GI:8347125
VERSION AF151972.1
KEYWORDS Mus musculus.
SOURCE Mus musculus.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

|||||
Db 453 CAGGGGATAGAGAGTGCATTAATCGTTAAVGGGAATCTAGAGACACAGCTCCAGACAC 512
QY 370 GAAGGATGCTCCCTCACCACCCAGCTCTTGAGGCAATGTGCACAGAGCCAGTCTGCACACC 429
Db 513 GAAGGAATGCTCCCTCACCACCCAGCTCTTGAGGCAATGTGCACAGAGCCAGTCTGCACACC 572
QY 430 AGAGACAGAGGCGCGAGGTCGAAGCTCCGGCTGTCTAAGAGGGAGGCTCCAGCCCTCT 489
Db 573 AGAGACAGAGGCGCGAGGTCGAAGCTCCGGCTGTCTAAGAGGGAGGCTCTCCAGCCCTCT 632
QY 490 GAATTACACGACGACATGACAGAGATGAGACAGAGATGATGAATGATGATGAGAA 549
Db 633 GAATTACACGACGACATGACAGAGATGAGACAGAGATGATGAATGATGATGAGAA 692
QY 550 TGGCTCGAATTTCTAATGCGCAAGCTCACCCGTGAGACCAAGACACAGAGAGCGCTC 609
Db 693 TGGCTCGAATTTCTAATGCGCAAGCTCACCCGTGAGACCAAGAGACACAGAGAGCGCTC 752
QY 610 TGAAGCGCGGCTGTCCGAACCGACATAGCAATGGGAACCTCCAGCTTGAGAGAGCAAG 669
Db 753 TGAAGCGCGGCTGTCCGAACCGACATAGCAATGGGAACCTCCAGCTTGAGAGAGCAAG 812
QY 670 AGCTCCCGCAGATCACCCGAGGTCCGACAGGCGCCACCATGTGACAGAGTACCTGT 729
Db 813 AGCTCCCGCAGATCACCCGAGGTCCGACAGGCGCCACCATGTGACAGAGTACCTGT 872
QY 730 GGAATTCGGGCTACACAGGTCTGGAGAGCTGCAGATCGTCTTACGCAAGACGCGCATG 789
Db 873 GGAATTCGGGCTACACAGGTCTGGAGAGCTGCAGATCGTCTTACGCAAGACGCGCATG 932
QY 790 GTCTCCCGCAGGCTGTCCGACCTCATGGAAGATGACACCTAAGAGCGCTACCTACCC 849
Db 933 GTCTCCCGCAGGCTGTCCGACCTCATGGAAGATGACACCTAAGAGCGCTACCTACCC 992
QY 850 ATCAGTTGACTTGAGCCAGAGTGGAGATCAGAGAGTGGATACACACAGAGTGGATGTC 909
Db 993 ATCAGTTGACTTGAGCCAGAGTGGAGATCAGAGAGTGGATACACACAGAGTGGATGTC 1052
QY 910 AGAGAGCAGAGATGAGACACAGACAGATATCAGAGATGATGAAGTTTGGATAGGTGA 969
Db 1053 AGAGAGCAGATGATGAGACACAGACAGATATCAGAGATGATGAAGTTTGGATAGGTGA 1112
QY 970 CCTGTGTGGGGAAGATCAAGGGCTCTCTGTGTGGCTGCGCATGGTGGTGTCTGGAA 1029
Db 1113 CCTGTGTGGGGAAGATCAAGGGCTCTCTGTGTGGCTGCGCATGGTGGTGTCTGGAA 1172
QY 1030 AGCCACCTCCAAAGCAGAGCCATGCCGGAATGCGCTGGGTACAGTGGTTGGTGAATG 1089
Db 1173 AGCCACCTCCAAAGCAGAGCCATGCCGGAATGCGCTGGGTACAGTGGTTGGTGAATG 1232
QY 1090 CAAGTTTCTGAGATCTCTGCTGACANAATGTGGTGGCTCTGAGCCGTTCAAGCAGCACTT 1149
Db 1233 CAAGTTTCTGAGATCTCTGCTGACANAATGTGGTGGCTCTGAGCCGTTCAAGCAGCACTT 1292
QY 1150 TAACTGGCTACCTCAATAAGCTGGTTCTTATAGGAAGGCCATGTAACACACTTCGGA 1209
Db 1293 TAACTGGCTACCTCAATAAGCTGGTTCTTATAGGAAGGCCATGTAACACACTTCGGA 1352
QY 1210 GAAAGCAGAGGTTGAGCTGSCAAGACCTTCTCCAGAGATGCTGGAAGTCACTGAGAGA 1259
Db 1353 GAAAGCAGAGGTTGAGCTGSCAAGACCTTCTCCAGAGATGCTGGAAGTCACTGAGAGA 1412
QY 1270 CCAAGCTAAGCCATGCTGAGTGGCCACAGGTGGCTTCAAGCCTACTGGATGAGAGG 1329
Db 1413 CCAAGCTAAGCCATGCTGAGTGGCCACAGGTGGCTTCAAGCCTACTGGATGAGAGG 1472
QY 1330 CCTCAAAACCAACAAGAGAACCAAGTGTAAATAGTCGAAGGTGGTGGTGTAGACAG 1389
Db 1473 CCTCAAAACCAACAAGAGAACCAAGTGTAAATAGTCGAAGGTGGTGGTGTAGACAG 1552
QY 1390 TAGGACTTAGAACCCAGAGAGCGGAGAACAAAAGTCSAAGAGCACAACCAATGATC 1449
|||||

Db 1533 TAGGAATTAAGAACCCAGAGAGCGGAGAACAAAAGTGAAGAGCACAACCAATGACTC 1592
QY 1450 TGTGCTTTGATGATCCCCCACCAGAGCCCTCTCAACAATAATAGTATATGGCGGAAGGA 1509
Db 1593 TGTGCTTTGATGATCCCCCACCAGAGCCCTCTCAACAATAATAGTATATGGCGGAAGGA 1652
QY 1510 CCGAGGGAGGATGAGAGAGAGAGCCGAGAACGATGAGCTTGAAGTCAACAAACAAGGG 1569
Db 1653 CCGAGGGAGGATGAGAGAGAGAGCCGAGAACGATGAGCTTGAAGTCAACAAACAAGGG 1712
QY 1570 CAATCTGAAGACCGCTGTTTGTCTGTGGAAGAAAGAACCTGTGTCTTCCACCCCT 1629
Db 1713 CAATCTGAAGACCGCTGTTTGTCTGTGGAAGAAAGAACCTGTGTCTTCCACCCCT 1772
QY 1630 CTTTGAAGGTGGCTCTGTGAGATTCGCCGAGTCCCTCTAGAGCTCTTTCATATGA 1689
Db 1773 CTTTGAAGGTGGCTCTGTGAGATTCGCCGAGTCCCTCTAGAGCTCTTTCATATGA 1832
QY 1690 TGATGAGGAGGCTATGATGCTTACCTGACCGTGTGCTGTGAGGGCCGTGACGTGCT 1749
Db 1833 TGATGAGGAGGCTATGATGCTTACCTGACCGTGTGCTGTGAGGGCCGTGACGTGCT 1892
QY 1750 GTGCACTAACAACAAGCTGTGACAGATGCTTGTGTGAGATGCTGTGAGAGTGTGTGG 1809
Db 1893 GTGCACTAACAACAAGCTGTGACAGATGCTTGTGTGAGATGCTGTGAGAGTGTGTGG 1952
QY 1810 CGCAGGCAACAGCTGAGATGCAAGCTGTGACAGAACCTGTGAGTGTGTATGTGCTGCC 1869
Db 1953 CGCAGGCAACAGCTGAGATGCAAGCTGTGACAGAACCTGTGAGTGTGTATGTGCTGCC 2012
QY 1870 TCAGCGCTGCATGGGTCCTCCGACGACAGAAAGATTGGAACTGGCGCTGCAAGACTT 1929
Db 2013 TCAGCGCTGCATGGGTCCTCCGACGACAGAAAGATTGGAACTGGCGCTGCAAGACTT 2072
QY 1930 CTTCACTAGTATCTTCACTGAGATTTGAGCAATTTGAGCCACCAAGTTTACCAAGCAATTC 1989
Db 2073 CTTCACTAGTATCTTCACTGAGATTTGAGCAATTTGAGCCACCAAGTTTACCAAGCAATTC 2132
QY 1990 TGCAGCCAAAAGAGGCCCATTTAGATCCCTGTCTCTTTGATGGAATTCGAACGGGGTA 2049
Db 2133 TGCAGCCAAAAGAGGCCCATTTAGATCCCTGTCTCTTTGATGGAATTCGAACGGGGTA 2192
QY 2050 CTTGTGCTCAAGAGATTTGGATTTAAGTGAAGAAATTAATGCTCCCAAGCTGTGCT 2109
Db 2193 CTTGTGCTCAAGAGATTTGGATTTAAGTGAAGAAATTAATGCTCCCAAGCTGTGCT 2252
QY 2110 AGAGTCCATGCGCTGTGGGAACCTGTTAAGCATGAAGGCCAGATCAATATGTCAATGACGT 2169
Db 2253 AGAGTCCATGCGCTGTGGGAACCTGTTAAGCATGAAGGCCAGATCAATATGTCAATGACGT 2312
QY 2170 CCGGAAAATACCAAGAAAATATTGAAGAGTGGGGCCGTTGCACTTGGTGGTGG 2229
Db 2313 CCGGAAAATACCAAGAAAATATTGAAGAGTGGGGCCGTTGCACTTGGTGGTGG 2372
QY 2230 AAGCCATGCAATATCTCTTACGTCATCAATCTGCGCCGAAAGAGTTTATATGAGGGCAC 2289
Db 2373 AAGCCATGCAATATCTCTTACGTCATCAATCTGCGCCGAAAGAGTTTATATGAGGGCAC 2432
QY 2290 AGGAAGGCTCTTCTTCAGAGTTTACCACTTGCTGAATTAACCGGCCCAAGAGAGGCGA 2349
Db 2433 AGGAAGGCTCTTCTTCAGAGTTTACCACTTGCTGAATTAACCGGCCCAAGAGAGGCGA 2492
QY 2350 CAACCGTCAATTTCTTGTGATGTTTGAAGATGTTGTGGCCATGAAATGAAATGACAAGAA 2409
Db 2493 CAACCGTCAATTTCTTGTGATGTTTGAAGATGTTGTGGCCATGAAATGAAATGACAAGAA 2552
QY 2410 AGACATGTCGAAGATTCTGCGATGTAACCCAGATGATGATGATGATGATGATGATGATGAT 2469
Db 2553 AGACATGTCGAAGATTCTGCGATGTAACCCAGATGATGATGATGATGATGATGATGATGAT 2612
QY 2470 TGTCTACAGGGCCCGGATCTTGTGGGTAACTTACCGGGAATGAAGAGCCCGTGAATGGC 2529
Db 2613 TGTCTACAGGGCCCGGATCTTGTGGGTAACTTACCGGGAATGAAG----- 2658

QY	2530	TTCAAGAAATGATAGCTCGAGCTGAGAGACTGCTGGAGTTGAGTAGACAGCAAAAGTT	2589
Db	2659	-----	2658
QY	2590	AAAGAAGTGCAGCAATTAACCAAGTCGAATCCATCAGACAGGGCAAAAACAGCT	2649
Db	2659	-----	2658
QY	2650	TTTCCCTGATGATGAATGCAAGAGACGATTTTGTGTGACATGAGCTCGAAGAGAT	2709
Db	2659	-----AGGAT-----	2663
QY	2710	CTTGGGCTTCCGTGCTACACAGGACGTCACCAATGAGGCGCGCGCGCTCAGAA	2769
Db	2664	CTTGGGCTTCCGTGCTACACAGGACGTCACCAATGAGGCGCGCGCGCTCAGAA	2723
QY	2770	GCTGCTGGGAGAGTCGCGAGAGTCGACGTCACGACACCTGTTTGGCCCTTGAAGA	2829
Db	2724	GCTGCTGGGAGAGTCGCGAGAGTCGACGTCACGACACCTGTTTGGCCCTTGAAGA	2783
QY	2830	CTACTTGGCTGTGAATAGTTCTTACCGAGACTGGGGAGCTGTGGGTACAGCCAGTGC	2889
Db	2784	CTACTTGGCTGTGAATAGTTCTTACCGAGAGTCGCGAGAGTCGCGTACAGCCAGTGC	2843
QY	2890	CAGAGTCACCCCTCCCTGAAGGACGTCACCTGTCCTTTTAACTCAGCTGTGGGG	2949
Db	2844	CAGAGTCACCCCTCCCTGAAGGACGTCACCTGTCCTTTTAACTCAGCTGTGGGG	2903
QY	2950	CCTCAGATCACTGATCACTGATCTTCTGCTGCTCAGTGGAGACAGCCCTTGGCCCTT	3009
Db	2904	CCTCAGATCACTGATCACTGATCTTCTGCTGCTCAGTGGAGACAGCCCTTGGCCCTT	2963
QY	3010	GCAGGGAGACCCCGGTGCTCCCTCGGTGTCACAGCTCAGACCTGCTGCTTAAAGTAGC	3069
Db	2964	GCAGGGAGACCCCGGTGCTCCCTCGGTGTCACAGCTCAGACCTGCTGCTTAAAGTAGC	3123
QY	3070	CCGGCAGTGGTGCATGTTCTTCTTACCTGAACCTTTAACTTAACTAGTAGTAGA	3129
Db	3024	CCGGCAGTGGTGCATGTTCTTCTTACCTGAACCTTTAACTTAACTAGTAGTAGA	3083
QY	3130	TGGCTTCTTAACTGCTGATGATTAATCACTGAGAGAGTGAAGTGAATACCAAAAA	3189
Db	3084	TGGCTTCTTAACTGCTGATGATTAATCACTGAGAGAGTGAAGTGAATACCAAAAA	3143
QY	3190	ACAAACAAAAACGAAACAAAAAACAACAAAAAAGCTCAACAGCTCTCTTGTACTCAGG	3249
Db	3144	ACAAACAAAAACGAAACAAAAAACAACAAAAAAGCTCAACAGCTCTCTTGTACTCAGG	3203
QY	3250	TTCAATGCTGCAAAATCACTGATGATTTTGTAAAGTAAACCGGTGCTCCACATTTGCTG	3309
Db	3204	TTCAATGCTGCAAAATCACTGATGATTTTGTAAAGTAAACCGGTGCTCCACATTTGCTG	3263
QY	3310	GAGATGCTATTTGTGATGATGCTGAGATGAGCAAGCTCAAGGGGCCAAAAAAATTC	3369
Db	3264	GAGATGCTATTTGTGATGATGCTGAGATGAGCAAGCTCAAGGGGCCAAAAAAATTC	3323
QY	3370	CCCTCTCCCCCGAGGATTTTGAAGATGATTTTATGTTTAACTTCTCCGACACTT	3429
Db	3324	CCCTCTCCCCCGAGGATTTTGAAGATGATTTTATGTTTAACTTCTCCGACACTT	3383
QY	3430	CCCTCTGCTTGTGATCAAGGCTGAGTCCGTTGGCTGCTGATTTTCCGAGAGAT	3489
Db	3384	CCCTCTGCTTGTGATCAAGGCTGAGTCCGTTGGCTGCTGATTTTCCGAGAGAT	3443
QY	3490	GATGTCAGCAGGATATACATGACACCTTTAGGCTTTTCCGAGAGGCGCCATGTGG	3549
Db	3444	GATGTCAGCAGGATATACATGACACCTTTAGGCTTTTCCGAGAGGCGCCATGTGG	3503
QY	3550	CTAGTCTCTCAGAGAGCTGAGTACATGTTTGGAGCTCAGGAAGGTGGCTGAGTGGC	3609
Db	3504	CTAGTCTCTCAGAGAGCTGAGTACATGTTTGGAGCTCAGGAAGGTGGCTGAGTGGC	3563
QY	3610	CTCTTCCAGGTGTGAGGATATGAGAGGAGAGCTTAGGAAATCCATTCCCACTCC	3669
Db	3564	CTCTTCCAGGTGTGAGGATATGAGAGGAGAGCTTAGGAAATCCATTCCCACTCC	3623
QY	3670	TCTTGCCAAATGAGGGGCCCATCCCAACAGCTCAGTCCCGAGAACCCCTTAGTCT	3729
Db	3624	TCTTGCCAAATGAGGGGCCCATCCCAACAGCTCAGTCCCGAGAACCCCTTAGTCT	3683
QY	3730	CATGAGAGCTAGACACAGACATGCTTCCCTTATCTGACAGTGTGGGAACT	3789
Db	3684	CATGAGAGCTAGACACAGACATGCTTCCCTTATCTGACAGTGTGGGAACT	3743
QY	3790	ACAGTGAACACCTTCTGAGATGTTAAAGCTTTTACCCACGATAGATTGTGTTT	3849
Db	3744	ACAGTGAACACCTTCTGAGATGTTAAAGCTTTTACCCACGATAGATTGTGTTT	3803
QY	3850	AGGGTCTTTTATGAGGGGATCAGTGCATGAGATGAAGAACCTGATTCAGAAATGCCAT	3909
Db	3804	AGGGTCTTTTATGAGGGGATCAGTGCATGAGATGAAGAACCTGATTCAGAAATGCCAT	3863
QY	3910	CGTATGCTTTTAAACACCTTTTACCTTATTAAGTGTATTTATAGAACAGACAA	3969
Db	3864	CGTATGCTTTTAAACACCTTTTACCTTATTAAGTGTATTTATAGAACAGACAA	3923
QY	3970	CACCTCTTTTATGACCTCAGACCTCTATTTTCAATGATTAACATTTTGTAACTGCG	4029
Db	3924	CACCTCTTTTATGACCTCAGACCTCTATTTTCAATGATTAACATTTTGTAACTGCG	3983
QY	4030	AGGTGTGGCTTTTGTAACTTACAGAGTGTGGGAGAGAGTGCCTGTTTCAAGCTT	4089
Db	3984	AGGTGTGGCTTTTGTAACTTACAGAGTGTGGGAGAGAGTGCCTGTTTCAAGCTT	4043
QY	4090	GTCCTCAGTGTCTTATTTTAACTTGTGCAAAAGATGACAGATGCCAGATTTACTTTC	4149
Db	4044	GTCCTCAGTGTCTTATTTTAACTTGTGCAAAAGATGACAGATGCCAGATTTACTTTC	4103
QY	4150	TGTTGATTTAAAGTGTGATTTCTCTTAAAAAATTTTTTTTTTTTTTTTTTTTTT	4195
Db	4104	TGTTGATTTAAAGTGTGATTTCTCTTAAAAAATTTTTTTTTTTTTTTTTTTTTT	4149
RESULT 9			
AF068628			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNML			
REMARK			
COMMENT			
FEATURES			
AF068628			
Mus musculus DNA cytosine-5 methyltransferase 3B3 (Dnm3b) mRNA,			
alternatively spliced, complete cds.			
AF068628.2 GI:6449473			
Mus musculus.			
Mus musculus.			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
1 (bases 1 to 3946)			
Okano, M., Xie, S. and Li, E.			
(Cytosine-5) methyltransferases			
Nat. Genet. 19 (3), 219-220 (1998)			
9662389			
2 (bases 1 to 3946)			
Xie, S., Okano, M. and Li, E.			
Direct Submision			
Submitted (28-MAY-1998) CVRC, Mass. Gen. Hospital, 149 13th Street,			
Charlestown, MA 02129, USA			
3 (bases 1 to 3946)			
Okano, M., Chijiwa, T., Sasaki, H. and Li, E.			
Direct Submision			
Submitted (04-NOV-1999) CVRC, Mass. Gen. Hospital, 149 13th Street,			
Charlestown, MA 02129, USA			
Sequence update by submitter			
On Nov 18 1999 this sequence version replaced gi:3327983.			
Location/Qualifiers			

source 1. .3946
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="2"
/map="84.0 cM"
gene 1. .3946
/gene="Dm13b"
misc-feature 1. .3946
/note="similar to EST sequences deposited in GenBank
Accession Numbers A116594, A119979, A117277, A1210568,
AA407106, and AA575617"
269. .2539
CDS
/gene="Dm13b"
/function="de novo DNA methylation"
/note="alternatively spliced product; contains Cys-rich
region"
/codon_start=1
/product="DNA cytosine-5 methyltransferase 3B3"
/protein_id="M449474"
/db_xref="GI:6449474"
/translation="MKGDSRLHNEEGASGYECITLVNNGFSQSSDPTKDAESPYLE
ALTEPVCPEPTEGRSSSLKREVSSSLNTYQDMTGDDRDDEVDGSDIIMPK
LRETRDTRRSBPARTRHSGTSLERORASPIITRQGRHHVOEYEPFAPTR
SRRRASSASPTWSSPASVDEMEVTPKSPSYDLSQDQEGMDITQVDAESRD
GDSLEYODDEKFGIDLVNKGIRKISPMWPMVSMKATSRQAMPGRWQWMDGDKF
SEISADKLVALGIFSOHPLNATPNKLYSRKAMYHLERARVAGTSSPCESLIED
QKPLHMAHAGFRKPTGISLAKPNKKOPEKRSRRRTTNDSSASESPPKLKTNSTGG
KORGEDESEKERNASEVTNNKMLEDRCLSCGKNPSEHLEGGIACOSCRDRLTEL
FYMYDDGQSYCTVCEBRELKLSNCSRCFCVECLVLAGAGTAEDAKLOEPLS
CYMCLPQRCGHJLRRRDMWMLRDEFTDPLEEFPKLYPAIDAKRPPRLVLSL
FDGIATGYLVLEKIKVEKYLAEVCAESIAVGTAKHEGOIIVRDKITKNTIEE
MGPFDIVIGSPCNDLSNVPARKGLIEGGRFLFPFYHLANTREKGNRPFFYMF
ENYNAKVNDKRISRLACNPYHIDAIKYSANRARIYPMGNLPGMRITGFPAHITD
VSMNGAROKLGRSVYFVIRHFLAPLDFACE"
BASE COUNT 983 a 1006 c 1056 g 901 t
ORIGIN
Query Match 87.7% Score 3677; DB 10; Length 3946;
Best Local Similarity 94.1% Pred. No. 0;
Matches 3946; Conservative 0; Mismatches 0; Indels 249; Gaps 2;
QY 1 GAATTCGGGGCGCGGGGTTAAGCGGCCCAAGTAACGTAGCGAGCGATCGCGCGGA 60
1 GAATTCGGGGCGCGGGGTTAAGCGGCCCAAGTAACGTAGCGAGCGATCGCGCGGA 60
Db 61 GATTCCGGAACCCGACACTCCGCGCGCGCGCGCGCGAGACCGCGCGCGATCGCGGC 120
61 GATTCCGGAACCCGACACTCCGCGCGCGCGCGCGCGAGACCGCGCGCGATCGCGGC 120
QY 121 GCGCGGTACAGCGAGCTCAGCAAGAGCGCGCTGTGTGCCAGACCTTGGAAAC 180
121 GCGCGGTACAGCGAGCTCAGCAAGAGCGCGCTGTGTGCCAGACCTTGGAAAC 180
Db 121 GCGCGGTACAGCGAGCTCAGCAAGAGCGCGCTGTGTGCCAGACCTTGGAAAC 180
121 GCGCGGTACAGCGAGCTCAGCAAGAGCGCGCTGTGTGCCAGACCTTGGAAAC 180
QY 181 TCAGGATATATACCTTTCAGACGCGGGATCTCCCTCCCTCCCATCATAGTGCCTTGGAC 240
181 TCAGGATATATACCTTTCAGACGCGGGATCTCCCTCCCTCCCATCATAGTGCCTTGGAC 240
Db 181 TCAGGATATATACCTTTCAGACGCGGGATCTCCCTCCCTCCCATCATAGTGCCTTGGAC 240
181 TCAGGATATATACCTTTCAGACGCGGGATCTCCCTCCCTCCCATCATAGTGCCTTGGAC 240
QY 241 CAATTCAGAGGCTCTCTTTCAGAAACAAATGAAGGAGACACGACATCTGAATGAAGA 300
241 CAATTCAGAGGCTCTCTTTCAGAAACAAATGAAGGAGACACGACATCTGAATGAAGA 300
Db 241 CAATTCAGAGGCTCTCTTTCAGAAACAAATGAAGGAGACACGACATCTGAATGAAGA 300
241 CAATTCAGAGGCTCTCTTTCAGAAACAAATGAAGGAGACACGACATCTGAATGAAGA 300
QY 301 AGAGGGTCCAGCGGGGATGAGAGAGTATATGCTTAATGGGAATCTTCACTAGTGCAGTC 360
301 AGAGGGTCCAGCGGGGATGAGAGAGTATATGCTTAATGGGAATCTTCACTAGTGCAGTC 360
Db 301 AGAGGGTCCAGCGGGGATGAGAGAGTATATGCTTAATGGGAATCTTCACTAGTGCAGTC 360
301 AGAGGGTCCAGCGGGGATGAGAGAGTATATGCTTAATGGGAATCTTCACTAGTGCAGTC 360
QY 361 CTCAGACACGAAAGATGCTCTCCACACCCAGCTTTCAGAGCAATCTGCAACAGAGCAGT 420
361 CTCAGACACGAAAGATGCTCTCCACACCCAGCTTTCAGAGCAATCTGCAACAGAGCAGT 420
Db 361 CTCAGACACGAAAGATGCTCTCCACACCCAGCTTTCAGAGCAATCTGCAACAGAGCAGT 420
361 CTCAGACACGAAAGATGCTCTCCACACCCAGCTTTCAGAGCAATCTGCAACAGAGCAGT 420
QY 421 CTCGACACCGAGAGACCGAGAGCGCGAGAGTCAAGCTCCCGGCTGTCTAAGAGGAGGTCTC 480
421 CTCGACACCGAGAGACCGAGAGCGCGAGAGTCAAGCTCCCGGCTGTCTAAGAGGAGGTCTC 480
Db 421 CTCGACACCGAGAGACCGAGAGCGCGAGAGTCAAGCTCCCGGCTGTCTAAGAGGAGGTCTC 480
421 CTCGACACCGAGAGACCGAGAGCGCGAGAGTCAAGCTCCCGGCTGTCTAAGAGGAGGTCTC 480

QY 481 CAGCCTTGTGAATTAACAGCAGAGCAATGACAGGAGTGGACAGAGATGATGAAGTGA 540
481 CAGCCTTGTGAATTAACAGCAGAGCAATGACAGGAGTGGACAGAGATGATGAAGTGA 540
Db 481 CAGCCTTGTGAATTAACAGCAGAGCAATGACAGGAGTGGACAGAGATGATGAAGTGA 540
481 CAGCCTTGTGAATTAACAGCAGAGCAATGACAGGAGTGGACAGAGATGATGAAGTGA 540
QY 541 TGAATGGGAATAGGCTGTGATATTCATGCAAAAGCTCACCCGTGAGACCAAGACACAG 600
541 TGAATGGGAATAGGCTGTGATATTCATGCAAAAGCTCACCCGTGAGACCAAGACACAG 600
Db 541 TGAATGGGAATAGGCTGTGATATTCATGCAAAAGCTCACCCGTGAGACCAAGACACAG 600
541 TGAATGGGAATAGGCTGTGATATTCATGCAAAAGCTCACCCGTGAGACCAAGACACAG 600
QY 601 GACGCGCTCGAAAGCCCGGCTCTCCGAAACCCGACATAGCATGSGACCTCCAGCTTGA 660
601 GACGCGCTCGAAAGCCCGGCTCTCCGAAACCCGACATAGCATGSGACCTCCAGCTTGA 660
Db 601 GACGCGCTCGAAAGCCCGGCTCTCCGAAACCCGACATAGCATGSGACCTCCAGCTTGA 660
601 GACGCGCTCGAAAGCCCGGCTCTCCGAAACCCGACATAGCATGSGACCTCCAGCTTGA 660
QY 661 GAGGCAAGAGCCCTCCCGCAAGTACACCGAGTGGCGAGGCGCGCACCATGTGACAGA 720
661 GAGGCAAGAGCCCTCCCGCAAGTACACCGAGTGGCGAGGCGCGCACCATGTGACAGA 720
Db 661 GAGGCAAGAGCCCTCCCGCAAGTACACCGAGTGGCGAGGCGCGCACCATGTGACAGA 720
661 GAGGCAAGAGCCCTCCCGCAAGTACACCGAGTGGCGAGGCGCGCACCATGTGACAGA 720
QY 721 GTACCCCTGTGAGATTCCGGCTACAGAGTCTCGAGACGTGAGCATCGCTTCCAGCAAG 780
721 GTACCCCTGTGAGATTCCGGCTACAGAGTCTCGAGACGTGAGCATCGCTTCCAGCAAG 780
Db 721 GTACCCCTGTGAGATTCCGGCTACAGAGTCTCGAGACGTGAGCATCGCTTCCAGCAAG 780
721 GTACCCCTGTGAGATTCCGGCTACAGAGTCTCGAGACGTGAGCATCGCTTCCAGCAAG 780
QY 781 CAGCCCATGTCATCCCTGCGACGCTGACCTTCATGGAAGAAGTACACCTTAAGAGCT 840
781 CAGCCCATGTCATCCCTGCGACGCTGACCTTCATGGAAGAAGTACACCTTAAGAGCT 840
Db 781 CAGCCCATGTCATCCCTGCGACGCTGACCTTCATGGAAGAAGTACACCTTAAGAGCT 840
781 CAGCCCATGTCATCCCTGCGACGCTGACCTTCATGGAAGAAGTACACCTTAAGAGCT 840
QY 841 CAGTACCCCATGTCATCCCTGCGACGCTGACCTTCATGGAAGAAGTACACCTTAAGAGCT 900
841 CAGTACCCCATGTCATCCCTGCGACGCTGACCTTCATGGAAGAAGTACACCTTAAGAGCT 900
Db 841 CAGTACCCCATGTCATCCCTGCGACGCTGACCTTCATGGAAGAAGTACACCTTAAGAGCT 900
841 CAGTACCCCATGTCATCCCTGCGACGCTGACCTTCATGGAAGAAGTACACCTTAAGAGCT 900
QY 901 GGTGATGTCAGAGAGCAGAGATGAGACACAGATATCAGAGATATGAAGAGTTTGG 960
901 GGTGATGTCAGAGAGCAGAGATGAGACACAGATATCAGAGATATGAAGAGTTTGG 960
Db 901 GGTGATGTCAGAGAGCAGAGATGAGACACAGATATCAGAGATATGAAGAGTTTGG 960
901 GGTGATGTCAGAGAGCAGAGATGAGACACAGATATCAGAGATATGAAGAGTTTGG 960
QY 961 AATAGGTGACCTGCTGTGGGAAAAGATCAAGGCTTCTCTGCTGGCTCCATGGTGGT 1020
961 AATAGGTGACCTGCTGTGGGAAAAGATCAAGGCTTCTCTGCTGGCTCCATGGTGGT 1020
Db 961 AATAGGTGACCTGCTGTGGGAAAAGATCAAGGCTTCTCTGCTGGCTCCATGGTGGT 1020
961 AATAGGTGACCTGCTGTGGGAAAAGATCAAGGCTTCTCTGCTGGCTCCATGGTGGT 1020
QY 1021 GTCTGGAAGACCAACCTCCAGGACAGGCGATGCGCGGATGCGGTACAGTGGTT 1080
1021 GTCTGGAAGACCAACCTCCAGGACAGGCGATGCGCGGATGCGGTACAGTGGTT 1080
Db 1021 GTCTGGAAGACCAACCTCCAGGACAGGCGATGCGCGGATGCGGTACAGTGGTT 1080
1021 GTCTGGAAGACCAACCTCCAGGACAGGCGATGCGCGGATGCGGTACAGTGGTT 1080
QY 1081 TGGTATGTCAGAGATTTTCTGAGATCTCTGTCGACAAATGCTGTGGGCTGTTCAG 1140
1081 TGGTATGTCAGAGATTTTCTGAGATCTCTGTCGACAAATGCTGTGGGCTGTTCAG 1140
Db 1081 TGGTATGTCAGAGATTTTCTGAGATCTCTGTCGACAAATGCTGTGGGCTGTTCAG 1140
1081 TGGTATGTCAGAGATTTTCTGAGATCTCTGTCGACAAATGCTGTGGGCTGTTCAG 1140
QY 1141 CCAAGCACTTAATCTGCTACCTTCAATTAAGCTGCTTCTTAATGAAGGCAATGACCA 1200
1141 CCAAGCACTTAATCTGCTACCTTCAATTAAGCTGCTTCTTAATGAAGGCAATGACCA 1200
Db 1141 CCAAGCACTTAATCTGCTACCTTCAATTAAGCTGCTTCTTAATGAAGGCAATGACCA 1200
1141 CCAAGCACTTAATCTGCTACCTTCAATTAAGCTGCTTCTTAATGAAGGCAATGACCA 1200
QY 1201 CACTGTGGAAGAACCGAGGCTGAGCTGGCAAGACCTTCCAGCAGTCTGAGAGATC 1260
1201 CACTGTGGAAGAACCGAGGCTGAGCTGGCAAGACCTTCCAGCAGTCTGAGAGATC 1260
Db 1201 CACTGTGGAAGAACCGAGGCTGAGCTGGCAAGACCTTCCAGCAGTCTGAGAGATC 1260
1201 CACTGTGGAAGAACCGAGGCTGAGCTGGCAAGACCTTCCAGCAGTCTGAGAGATC 1260
QY 1261 ACTGGAGACCAAGTGAAGCCATCTGAGAGTGGGCCACAGGTGGCTTCAAGGCTTAC 1320
1261 ACTGGAGACCAAGTGAAGCCATCTGAGAGTGGGCCACAGGTGGCTTCAAGGCTTAC 1320
Db 1261 ACTGGAGACCAAGTGAAGCCATCTGAGAGTGGGCCACAGGTGGCTTCAAGGCTTAC 1320
1261 ACTGGAGACCAAGTGAAGCCATCTGAGAGTGGGCCACAGGTGGCTTCAAGGCTTAC 1320
QY 1321 GATCGAGGCGCTCAAAACCAAGAAAGACCAAGAGTGTAAATGATGCAAGGTGCTCG 1380
1321 GATCGAGGCGCTCAAAACCAAGAAAGACCAAGAGTGTAAATGATGCAAGGTGCTCG 1380
Db 1321 GATCGAGGCGCTCAAAACCAAGAAAGACCAAGAGTGTAAATGATGCAAGGTGCTCG 1380
1321 GATCGAGGCGCTCAAAACCAAGAAAGACCAAGAGTGTAAATGATGCAAGGTGCTCG 1380
QY 1381 TTCAGACAGTAGGAATTAACCCAGAGAACCGAGAGACAAAAGTTCGAGAGCACAAC 1440
1381 TTCAGACAGTAGGAATTAACCCAGAGAACCGAGAGACAAAAGTTCGAGAGCACAAC 1440
Db 1381 TTCAGACAGTAGGAATTAACCCAGAGAACCGAGAGACAAAAGTTCGAGAGCACAAC 1440
1381 TTCAGACAGTAGGAATTAACCCAGAGAACCGAGAGACAAAAGTTCGAGAGCACAAC 1440
QY 1441 CAATGACTGTGCTGCTTCTGAGTCCCGCCACCCAGCGCTCAAGCAAAATGCTATGG 1500
1441 CAATGACTGTGCTGCTTCTGAGTCCCGCCACCCAGCGCTCAAGCAAAATGCTATGG 1500
Db 1441 CAATGACTGTGCTGCTTCTGAGTCCCGCCACCCAGCGCTCAAGCAAAATGCTATGG 1500
1441 CAATGACTGTGCTGCTTCTGAGTCCCGCCACCCAGCGCTCAAGCAAAATGCTATGG 1500
QY 1501 CGGGAAGAGACCGAGGAGAGATGAGAGACCCGAGAACGATGCTTCTAATTCACCAA 1560
1501 CGGGAAGAGACCGAGGAGAGATGAGAGACCCGAGAACGATGCTTCTAATTCACCAA 1560
Db 1501 CGGGAAGAGACCGAGGAGAGATGAGAGACCCGAGAACGATGCTTCTAATTCACCAA 1560
1501 CGGGAAGAGACCGAGGAGAGATGAGAGACCCGAGAACGATGCTTCTAATTCACCAA 1560

QY	1561	CAACAAAGGCAAATCTGGAAGACCGCGCTGTTTGTCTCTGTGTGGAAGAAGAACCCCTGTCTCTT	1620
Db	1501	CAACAAAGGCAATCTGGAAGACCGCGCTGTTTGTCTCTGTGTGGAAGAAGAACCCCTGTCTCTT	1560
QY	1621	CCACCCCCCTTTTGGAGGGTGGGGCTCTGTGCAGAGTTTCCCGGGATCGCTCTCTTAGAGCTCTT	1680
Db	1561	CCACCCCCCTTTTGGAGGGTGGGGCTCTGTGTGAGAGTTTCCCGGGATCGCTCTCTTAGAGCTCTT	1620
QY	1681	CTACATCTATGATGAGAGACGGCTATACAGTCTCTACTGACACCGCTGTCTGTGAGAGCCGTGA	1740
Db	1621	CTACATCTATGATGAGAGACGGCTATACAGTCTCTACTGACACCGCTGTCTGTGAGAGCCGTGA	1680
QY	1741	ACTGCTCTCTGTGAGTAACACAAAGCTCTGTGCAGATGCTCTGTGTGTGAGTGTCTGAGAGT	1800
Db	1681	ACTGCTCTCTGTGAGTAACACAAAGCTCTGTGCAGATGCTCTGTGTGTGAGTGTCTGAGAGT	1740
QY	1801	GCTGGTGGGGCGCAGGACAGCTGAGATGTGCACCAAGCTGGCAGGAAACCCCTGAGCTGTATAT	1860
Db	1741	GCTGGTGGGGCGCAGGACAGCTGAGATGTGCACCAAGCTGGCAGGAAACCCCTGAGCTGTATAT	1800
QY	1861	GTGGCTCCCTCAGCGGTGCATGTGGGGTCCCGACAGCAGGAAACATTTGGAACATCGCGCT	1920
Db	1801	GTGGCTCCCTCAGCGGTGCATGTGGGGTCCCGACAGCAGGAAACATTTGGAACATCGCGCT	1860
QY	1921	GGAAAGCTCTTCTACACTAGATCCTGTACGCTGGAGAAATTTAGGACACCCAGCTGTACC	1980
Db	1861	GGAAAGCTCTTCTACACTAGATCCTGTACGCTGGAGAAATTTAGGACACCCAGCTGTACC	1920
QY	1981	AGCAATTCCTGTGACGCCAAAGAGAGGCCCAATTAGAGTCCGTCTGTGTGATGGAATATGC	2040
Db	1921	AGCAATTCCTGTGACGCCAAAGAGAGGCCCAATTAGAGTCCGTCTGTGTGATGGAATATGC	1980
QY	2041	AACGGGGTACTGTGTCTCAGAGAGTTGGGTATTTAAAGTGGAAAGTACATTTGCTCCGA	2100
Db	1981	AACGGGGTACTGTGTCTCAGAGAGTTGGGTATTTAAAGTGGAAAGTACATTTGCTCCGA	2040
QY	2101	AGCTGTGTGCAGAGTCCATCGCTGTGTGGAGCTGTTAAGCTAGAGGCCAGATCAAAATATGT	2160
Db	2041	AGCTGTGTGCAGAGTCCATCGCTGTGTGGAGCTGTTAAGCTAGAGGCCAGATCAAAATATGT	2100
QY	2161	CAATGACGTCGGAAAAATCACCAGAAAAAATATTGGAAGTGGGGCCCGTTGCACTGGT	2220
Db	2101	CAATGACGTCGGAAAAATCACCAGAAAAAATATTGGAAGTGGGGCCCGTTGCACTGGT	2160
QY	2221	GATTGTGTGAAGCCCATGTCAATGATCTCTTAACGTCATTCCTGCCCGAAAGTTTATA	2280
Db	2161	GATTGTGTGAAGCCCATGTCAATGATCTCTTAACGTCATTCCTGCCCGAAAGTTTATA	2220
QY	2281	TGAGGGGCACAGAAAGGCTTTCTTTCGAGATTTTACCACTTGTGTAATTTACCCGCCCA	2340
Db	2221	TGAGGGGCACAGAAAGGCTTTCTTTCGAGATTTTACCACTTGTGTAATTTACCCGCCCA	2280
QY	2341	GGAGGGGCACACCGTCACATTTCTTGTGAGTTTCCAGATGTGGCCATGTAAATGTA	2400
Db	2281	GGAGGGGCACACCGTCACATTTCTTGTGAGTTTCCAGATGTGGCCATGTAAATGTA	2340
QY	2401	TGACAAAGAAACACATCTCAAGATTCCTGTGGCATGTAAACCAAGATGATCATCCATCAA	2460
Db	2341	TGACAAAGAAACACATCTCAAGATTCCTGTGGCATGTAAACCAAGATGATCATCCATCAA	2400
QY	2461	GGGTGTCCTGTCTACAGAGGCCCGGTACTTCTGTGGGTAACTACCCGAATGAAGAGCGC	2520
Db	2401	GGGTGTCCTGTCTACAGAGGCCCGGTACTTCTGTGGGTAACTACCCGAATGAAGAGCGC	2460
QY	2521	CGTGATGCTTCAAGATGATTAAGCTGTGAGCTGTGAGAGCTCCTCGAGTTCAAGTAGAGAC	2580
Db	2461	CGTGATGCTTCAAGATGATTAAGCTGTGAGCTGTGAGAGCTCCTCGAGTTCAAGTAGAGAC	2520
QY	2581	AGCAAAAGTTAAAGAAAGTCAGACAATACCAACCAAGTCGAACTCATACAGAGGCA	2640
Db	2521	AGCAAAAGTTAAAGAAAGTCAGACAATACCAACCAAGTCGAACTCATACAGAGGCA	2580
QY	2641	AAACCAAGCTTTCCCTGTATGTCATTAATGCAAGAGACGAGCTTTTGTGTGTGACATGAGCT	2700

Db	2456	-----	2455
OY	2701	CGAAGGATCTTCGGCTTCCTGCTCATACACGAGCTGTCCACATAGGCCGCCGCCG	2760
Db	2456	----AGATCTTCGGCTTCCTGCTCATACACGAGCTGTCCACATAGGCCGCCGCCG	2511
OY	2761	CCGTGCAAGAGCTGCTGGGCGAGTCTCGGAGTGTACCGGTATCAGACACTGTGTGGCCC	2820
Db	2512	CCGTGCAAGAGCTGCTGGGCGAGTCTCGGAGTGTACCGGTATCAGACACTGTGTGGCCC	2571
OY	2821	CTTGAAAGACTCTTGTGCTGTGAATAGTTTACCCAGAGACTGGGAGTCTCGGTGAGA	2880
Db	2572	CTTGAAAGACTCTTGTGCTGTGAATAGTTTACCCAGAGACTGGGAGTCTCGGTGAGA	2631
OY	2881	GGCAGTGGCCAGAGTACACCCCTGCTCGAAGGAGACTCACCGTCCCTTTTACGTACCC	2940
Db	2632	GGCAGTGGCCAGAGTACACCCCTGCTCGAAGGAGACTCACCGTCCCTTTTACGTACCC	2691
OY	2941	TGTGTGGGGCTCAATATCATCTGTACCTCAGTTCCTCTGTGATAGTGGAGAGAGCTC	3000
Db	2692	TGTGTGGGGCTCAATATCATCTGTACCTCAGTTCCTCTGTGATAGTGGAGAGAGCTC	2751
OY	3001	CTGGCCCTTCGAGGGAGACCCCGGTGCTCCCTGTGTGCACAGTCAACCTGGCTGT	3060
Db	2752	CTGGCCCTTCGAGGGAGACCCCGGTGCTCCCTGTGTGCACAGTCAACCTGGCTGT	2811
OY	3061	TAGAGTATGCCCGCATGTGTGCTCATGTCTCTTACCTCGAAACTTTAAACTTTCGAATG	3120
Db	2812	TAGAGTATGCCCGCATGTGTGCTCATGTCTCTTACCTCGAAACTTTAAACTTTCGAATG	2871
OY	3121	GTATATAGATGGCTTCTTTTACCCTCTGAGTTTATCACTCAGAGTAGTGTCTAAGAT	3180
Db	2872	GTATATAGATGGCTTCTTTTACCCTCTGAGTTTATCACTCAGAGTAGTGTCTAAGAT	2931
OY	3181	ACCAAAAAAACAACAAAAACAGAAACAAAAAACAACAAAAAACCCTACACAGCTCTCTTA	3240
Db	2932	ACCAAAAAAACAACAAAAACAGAAACAAAAAACAACAAAAAACCCTACACAGCTCTCTTA	2991
OY	3241	GTACACAGTTCCTGTGTGCAAAATCACTTGTAGTTTGTTTTTTAAGTAACCCGTCTCCA	3300
Db	2992	GTACACAGTTCCTGTGTGCAAAATCACTTGTAGTTTGTTTTTTAAGTAACCCGTCTCCA	3051
OY	3301	CATTTCGTGGAGGATCTATTGTGAATGTGGGTGAGATGAGCAAGTCAAGGGCCAAA	3360
Db	3052	CATTTCGTGGAGGATCTATTGTGAATGTGGGTGAGATGAGCAAGTCAAGGGCCAAA	3111
OY	3361	AAAAATTCCTCCCTCTCCGCCACGAGATTTTGAAGATGATTTATGTGTTAAGTCTCC	3420
Db	3112	AAAAATTCCTCCCTCTCCGCCACGAGATTTTGAAGATGATTTATGTGTTAAGTCTCC	3171
OY	3421	TGGCACCTTCCCTCTCTTGGTATACAAAGGCTGAAAGTCTGTGGTGTATGACATTTCC	3480
Db	3172	TGGCACCTTCCCTCTCTTGGTATACAAAGGCTGAAAGTCTGTGGTGTATGACATTTCC	3231
OY	3481	CAGATGATATATGTACGACGAGGATACATCCACTTTAGGCTTTTCCCTGGCAGGGG	3540
Db	3232	CAGATGATATATGTACGACGAGGATACATCCACTTTAGGCTTTTCCCTGGCAGGGG	3291
OY	3541	CCCATGTGGCTAGTCTCTCACGAAGACTGGAGTGAATGTTTGGAGCTCAGGAAGGGTGG	3600
Db	3292	CCCATGTGGCTAGTCTCTCACGAAGACTGGAGTGAATGTTTGGAGCTCAGGAAGGGTGG	3351
OY	3601	TGGATGTGCCCTCTTCACAGTGTGAGGGAATACGAAGAGAGAAAGCTTAGGAAATCCATT	3660
Db	3352	TGGATGTGCCCTCTTCACAGTGTGAGGGAATACGAAGAGAGAAAGCTTAGGAAATCCATT	3411
OY	3661	CCCATGCTCCTTGGCCAAATGAGGGGCCAGTCCCAACAGCTAGGTCCCCAGAAACCC	3720
Db	3412	CCCATGCTCCTTGGCCAAATGAGGGGCCAGTCCCAACAGCTAGGTCCCCAGAAACCC	3471
OY	3721	CTAAGTCTCATGAGAACTAGAGACGAAGAGCAATCTGTCCTCTTATCGACAGATGT	3780

Db 3472 CTAGTTCCTCATGAGAACCTAGACACAGACACATCTCCCTTATCTAGACAGTGT 3531
QY 3781 TGGGGAAGTACAGTGAACCTCTGAGATGTTAAAGCTTTTACCCCAAGATGTT 3840
Db 3532 TGGGGAAGTACAGTGAACCTCTGAGATGTTAAAGCTTTTACCCCAAGATGTT 3591
QY 3841 GTGTTTAAAGGGGTCTTTTAAAGGGGATCATGTGAGATGAAGAAGATGATTTAG 3900
Db 3592 GTGTTTAAAGGGGTCTTTTAAAGGGGATCATGTGAGATGAAGAAGATGATTTAG 3651
QY 3901 AAATGCCATGTATGTTTAAACACCTTTTACCTAATTAACAGTCTAATTTATAGA 3960
Db 3652 AAATGCCATGTATGTTTAAACACCTTTTACCTAATTAACAGTCTAATTTATAGA 3711
QY 3961 AGCAGACACACTCTTTTAAAGTCTGACATCTCTAATTTACATGATTTATTTT 4020
Db 3712 AGCAGACACACTCTTTTAAAGTCTGACATCTCTAATTTACATGATTTATTTT 3771
QY 4021 GTAATCCGCAAGTGTGGCTTTTGTAACTTACAGTGTGGGAGAGACTGCTTGT 4080
Db 3772 GTAATCCGCAAGTGTGGCTTTTGTAACTTACAGTGTGGGAGAGACTGCTTGT 3831
QY 4081 CAACAGTTTGTCTCAGTGTGTTTCTAATTTTAAAGTGAAGATGACAGATGCCAGAT 4140
Db 3832 CAACAGTTTGTCTCAGTGTGTTTCTAATTTTAAAGTGAAGATGACAGATGCCAGAT 3891
QY 4141 TTACTTCTGTTGATTAAGTGTATTTCTCTTAAATTTTAAATTTTAAATTTTAAAT 4195
Db 3892 TTACTTCTGTTGATTAAGTGTATTTCTCTTAAATTTTAAATTTTAAATTTTAAAT 3946

RESULT 10 4089 bp mRNA linear ROD 08-JUN-2000
AF151971
LOCUS Mus musculus DNA cytosine-specific methyltransferase isoform 3
DEFINITION (Dnmt3b) mRNA, complete cds.
ACCESSION AF151971
VERSION AF151971.1 GI:8347122
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 4089)
AUTHORS Yiu, B., Chen, Y.T., Zhu, M., Luo, Y.J., Zhu, N., Xu, S.C., Wu, G.Y. and
Shen, Y.
TITLE Cloning of full-length Dnmt3b cDNA and its alternative splicing
isoforms in mouse embryonic tissue
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 4089)
Yiu, B., Chen, Y.T., Zhu, M., Luo, Y.J., Zhu, N., Xu, S.C., Wu, G.Y. and
Shen, Y.
TITLE Direct Submission
JOURNAL Submitted (17-MAY-1999) Department of Biochemistry and Molecular
Biology, Institution of Basic Medical Sciences, Chinese Academy of
Medical Sciences, Dong Dan San Tiao 5, Beijing City 100005,
P.R.China
FEATURES
source 1. 4089
location/Qualifiers
1. 4089
/organism="Mus musculus"
/strain="KM"
/db_xref="taxon:10090"
/dev_stage="8-9 day old embryo"
1. 4089
/gene="Dnmt3b"
412..2742
/gene="Dnmt3b"
/function="methylates cytosine in DNA"
/note="alternatively spliced"
/codon_start=1
/product="DNA cytosine-specific methyltransferase isoform
3"
/protein_id="AAAF74517.1"

FEATURES
source 1. 4089
location/Qualifiers
1. 4089
/organism="Mus musculus"
/strain="KM"
/db_xref="taxon:10090"
/dev_stage="8-9 day old embryo"
1. 4089
/gene="Dnmt3b"
412..2742
/gene="Dnmt3b"
/function="methylates cytosine in DNA"
/note="alternatively spliced"
/codon_start=1
/product="DNA cytosine-specific methyltransferase isoform
3"
/protein_id="AAAF74517.1"

/db_xref="GI:8347123"
/translation="MKGDSRLHLEDEAGSYEECLIVNGNFSDSDSPDKAPSPVLE
AICREPTCEPTRRGRSSRLSKREVSLNLTVDNMKGDRDSDSDVDDGNSDILMK
LIRETKDTRTSESPAVTRHSNKTSSLIERASPRITRGRGHHVOEYVEYEPATR
SRRRRASSASTPWSVPASVDFMEVEYKSVSTVSVDLSQDIDEGMDPTVOMDESLY
GDSTYEODDKFEFGIDLVWYKIKFSNMPAVVWMTKSKROADMGMVOMFDDGK
SEISADKLVALGFSQHNLTAPKLVASRYAMHTLEKARVAGKTFSSSPGSLSD
OLKPLMEAHAGGKFPDTGEGKLPKPKOENKSRRTTDSASRSPPEKRLKTFNYSYG
KDRDEDESREPMASVYTNMKNLIEDRLSGKNRPVSPHPLFPGGLCSGDRDELFL
FYRDEDEIYSTCTVCCBRELILCSNTSCRCRCVBELEVLYVAGTAEDKLDPEPS
CYMCLPQRCHGVLRNRDWMRLDFTTDDLEFEPKLYPAIPAKRPIVLSL
FDGIATGTLVLEKELIYKVEYIASEVASEIAVGTVHEGQIKYVNDKRTKKNIEE
MGPDVLVGGSPCNDLSVNPARKLIEGSLRLEFFYHLNLTFRPEKGNRPFPMF
ENVVAMVNDKKDISRFLACPVMDIAIKVSAARVAFMNLGMNRIRGFPAHYTD
VSNMGRCAKOLLRSMSVPIRRLFLPNDYFACE"
BASE COUNT 995 a 1071 c 1104 g 919 t
ORIGIN
Query Match 87.4%; Score 3664.8; DB 10; Length 4089;
Best Local Similarity 94.0%; Pred. No. 0;
Matches 3935; Conservative 0; Mismatches 2; Indels 249; Gaps 2;
Db 10 GCGCCGGGTTAAGCGCCCAAGTAACGACGACGATGCGCGCCGAGATTCGCGA 69
QY 153 GCGCCGGGTTAAGCGCCCAAGTAACGACGACGATGCGCGCCGAGATTCGCGA 212
Db 70 ACCGACACTCCGCTA 129
QY 213 ACCGACACTCCGCTA 272
Db 130 CAGCAGCGCTCAGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTA 189
QY 273 CAGCAGCGCTCAGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTA 332
Db 190 TACCTTCCAGACGCGGATCTCCCTCCCTCCATCATATGTCCTTGGACCAATTCAG 249
QY 333 TACCTTCCAGACGCGGATCTCCCTCCCTCCATCATATGTCCTTGGACCAATTCAG 392
Db 250 GGCCTCTTTCAGGAAACATGAAGGAGACAGCAGCATGTAATGAAGAAGGCGTC 309
QY 393 GGCCTCTTTCAGGAAACATGAAGGAGACAGCAGCATGTAATGAAGAAGGCGTC 452
Db 310 CAGCGGATGAGAGAGTGCATTAATGGAAGTCAAGTGCAGAGTCTCAGACAC 369
QY 433 CAGCGGATGAGAGAGTGCATTAATGGAAGTCAAGTGCAGAGTCTCAGACAC 512
Db 370 GAAGATGCTCCCTCAGCCCACTTCTGAGGCAATGTCACAGACGATGTCACACC 429
QY 513 GAAGATGCTCCCTCAGCCCACTTCTGAGGCAATGTCACAGACGATGTCACACC 572
Db 430 AGAGACAGAGCGCGCAGTCAAGTCCCGGTCTTAAGAGGAGGTCCTCAGGCTCT 489
QY 573 AGAGACAGAGCGCGCAGTCAAGTCCCGGTCTTAAGAGGAGGTCCTCAGGCTCT 632
Db 490 GAATTAACGCGCAGCAGATGAGAGATGAGAGACAGATGATGAAGTATGATGAGAA 549
QY 633 GAATTAACGCGCAGCAGATGAGAGATGAGAGACAGATGATGAAGTATGATGAGAA 692
Db 550 TGGCTGTATATTTCTAATGCCAAGCTCAGCCGCTGAGACCAAGACACAGAGCGGTC 609
QY 633 TGGCTGTATATTTCTAATGCCAAGCTCAGCCGCTGAGACCAAGACACAGAGCGGTC 752
Db 610 TGAAGCCCGGTGTGCGCAACCGCATGAGCAATGGAGACCTCCACTTGGAGAGGCAAG 669
QY 753 TGAAGCCCGGTGTGCGCAACCGCATGAGCAATGGAGACCTCCACTTGGAGAGGCAAG 812
Db 670 AGCCTCCCGCAGAAATCACCCGAGGTGCGCAGAGGCGCGCAGCATGTCAGAGATACCTGT 729
QY 813 AGCCTCCCGCAGAAATCACCCGAGGTGCGCAGAGGCGCGCAGCATGTCAGAGATACCTGT 872
Db 730 GGAATTTCCGCTACAGAGTCTGAGAGAGCTGTCAGATGTCCTTACGACAGCCGATG 789
QY 873 GGAATTTCCGCTACAGAGTCTGAGAGAGCTGTCAGATGTCCTTACGACAGCCGATG 932

QY 790 GTCATCCCTGCGACCGTGCATTCATGGAAGTGAACCTAGAGCGCTCACTACCCC 849
DB 933 GTGATCCCTGCGACCGTGCATTCATGGAAGTGAACCTAGAGCGCTCACTACCCC 992
QY 850 ATCAGTGGACTGAGCCAGATGAGATCAGAGGATGATGATACCAACAGTGTATGC 909
DB 993 ATCAGTGGACTGAGCCAGATGAGATCAGAGGATGATGATACCAACAGTGTATGC 1052
QY 910 AGAGCAGATGAGAGACACAGAGATATCAGATGATTAAGAGTTTGAATAGTGA 969
DB 1053 AGAGCAGATGAGAGACACAGAGATATCAGATGATTAAGAGTTTGAATAGTGA 1112
QY 970 CCTGCTGGGGAAGATCAAGGGCTTCCTGCTGGGCTGCTGATGATGATGATGATG 1029
DB 1113 CCTGCTGGGGAAGATCAAGGGCTTCCTGCTGGGCTGCTGATGATGATGATGATG 1172
QY 1030 AGCCACTCCAGAGCAGAGCGCATGCGGGAATGCGGTGATGATGATGATGATG 1089
DB 1173 AGCCACTCCAGAGCAGAGCGCATGCGGGAATGCGGTGATGATGATGATGATG 1232
QY 1090 CAAGTTTCTGAGATCTCTGCTGACAACTGTGCTCTGGGCTGCTGACCTGACT 1149
DB 1233 CAAGTTTCTGAGATCTCTGCTGACAACTGTGCTCTGGGCTGCTGACCTGACT 1292
QY 1150 TAATCTGGCTACCTTCAATTAAGCTGTTCTTATAGGAAGCCATGATACCACTGGA 1209
DB 1293 TAATCTGGCTACCTTCAATTAAGCTGTTCTTATAGGAAGCCATGATACCACTGGA 1352
QY 1210 GAAAGCCAGGTTTCAGAGTGGCAAGACCTTCTCAGACAGTCTCTGAGAGTCACTGAGGA 1269
DB 1353 GAAAGCCAGGTTTCAGAGTGGCAAGACCTTCTCAGACAGTCTCTGAGAGTCACTGAGGA 1412
QY 1270 CCAGCTGAAGCCATGCTGAGTGGGCGCAAGGCTGCTCAAGCTACTGAGATGAGG 1329
DB 1413 CCAGCTGAAGCCATGCTGAGTGGGCGCAAGGCTGCTCAAGCTACTGAGATGAGG 1472
QY 1330 CCTCAAAACCAACAGAGCAACAGTGTATTAAGTGAAGTGTGCTGCTGAGAG 1389
DB 1473 CCTCAAAACCAACAGAGCAACAGTGTATTAAGTGAAGTGTGCTGCTGAGAG 1497
QY 1390 TAGGAACCTTGAACCCAGAGAGCGAGAACAAAGTGCAGAGCAGACACATGACTC 1449
DB 1498 TAGGAACCTTGAACCCAGAGAGCGAGAACAAAGTGCAGAGCAGACACATGACTC 1532
QY 1450 TGTGCTTCTGAGTCCCCCCCCCAAGCGCTCAAGAAATAGCTATGGGGAGGA 1509
DB 1533 TGTGCTTCTGAGTCCCCCCCCCAAGCGCTCAAGAAATAGCTATGGGGAGGA 1592
QY 1510 CCGAGGGGAGATGAGAGAGCCGAGAAAGGATGCTTGAAGTACCAACACAGAGG 1569
DB 1593 CCGAGGGGAGATGAGAGAGCCGAGAAAGGATGCTTGAAGTACCAACACAGAGG 1652
QY 1570 CAATCTGGAAGACCGCTGTTCTGTGGAAGAGAACCCGTGTCTTCACCCCT 1629
DB 1653 CAATCTGGAAGACCGCTGTTCTGTGGAAGAGAACCCGTGTCTTCACCCCT 1712
QY 1630 CTTTGAGGGTGGCTCTGTCAGAGTGGCGGATGCTTCTAGAGCTTCTTACATSTA 1689
DB 1713 CTTTGAGGGTGGCTCTGTCAGAGTGGCGGATGCTTCTAGAGCTTCTTACATSTA 1772
QY 1690 TGATGAGAGGAGGCTATCACTGACAGCGGTGTGTAAGGGCCGTGAACCTCTCT 1749
DB 1773 TGATGAGAGGAGGCTATCACTGACAGCGGTGTGTAAGGGCCGTGAACCTCTCT 1832
QY 1750 GTGCAATACCAAGCTGTCAGAGTGGCTGATGCTGTCAGAGGCGGAGTCTCTCT 1809
DB 1833 GTGCAATACCAAGCTGTCAGAGTGGCTGATGCTGTCAGAGGCGGAGTCTCTCT 1892
QY 1810 GCGAGGCAAGCTGAGAGTCCAGAGTCCAGAGACCTGAGACTGCTATATGTGCTCC 1869
DB 1893 GCGAGGCAAGCTGAGAGTCCAGAGTCCAGAGACCTGAGACTGCTATATGTGCTCC 1952
QY 1870 TCAGCGCTGCCATGGGTTCTCCGAGCGAGGAAGATTTGAGACATGCGCTGACACTT 1929
DB 1953 TCAGCGCTGCCATGGGTTCTCCGAGCGAGGAAGATTTGAGACATGCGCTGACACTT 2012
QY 1930 CTTCACTACTGATCTGATGCTGAGAGATTTTGGCCACCAAGTGTATGAGCAATTC 1989
DB 2013 CTTCACTACTGATCTGATGCTGAGAGATTTTGGCCACCAAGTGTATGAGCAATTC 2072
QY 1990 TGCAGCCAAAAGAGGCCATTAGACTGCTCTGTTGATGATGATTTGCAAGCGGTA 2049
DB 2073 TGCAGCCAAAAGAGGCCATTAGACTGCTCTGTTGATGATGATTTGCAAGCGGTA 2132
QY 2050 CTTGGTCTCAGAGAGTTGGGTATTAAGTGAAGAAAGTACATTCCTCGAAGTGTGC 2109
DB 2133 CTTGGTCTCAGAGAGTTGGGTATTAAGTGAAGAAAGTACATTCCTCGAAGTGTGC 2192
QY 2110 AGAGTCAATGCTGTGGGAAGTGTAAAGCATGAAAGCCAGATCAATATGTCAATGAGCT 2169
DB 2193 AGAGTCAATGCTGTGGGAAGTGTAAAGCATGAAAGCCAGATCAATATGTCAATGAGCT 2252
QY 2170 CCGGAAATACCAAGAAATATTAAGAGTGGGGCCGTTGACTTGGTATGATG 2229
DB 2253 CCGGAAATACCAAGAAATATTAAGAGTGGGGCCGTTGACTTGGTATGATG 2312
QY 2230 AACCCATGCAATGATCTCTTAAGTCAATGCTGCGGCAAGGTTTATATGAGGCGAC 2289
DB 2313 AACCCATGCAATGATCTCTTAAGTCAATGCTGCGGCAAGGTTTATATGAGGCGAC 2372
QY 2290 AGAAGGCTCTTCTGAGATTTTACAGTGTGTAATTAACCGGCGGCAAGAGGCGCA 2349
DB 2373 AGAAGGCTCTTCTGAGATTTTACAGTGTGTAATTAACCGGCGGCAAGAGGCGCA 2432
QY 2350 CAACCGTCAATCTTCTGATGATGTCAGAAATGTTGGCCATGAATGAATGAGAGGA 2409
DB 2433 CAACCGTCAATCTTCTGATGATGTCAGAAATGTTGGCCATGAATGAATGAGAGGA 2492
QY 2410 AGACATCTCAAGATTCCTGAGATGTAACCCAGATGATGATGATGATGATGATG 2469
DB 2493 AGACATCTCAAGATTCCTGAGATGTAACCCAGATGATGATGATGATGATGATG 2552
QY 2470 TGTCTACAGGGCGGCTGCTGAGTGGGTTAACTACCGGAAATGAACGGCGCTGATG 2529
DB 2553 TGTCTACAGGGCGGCTGAGTGGGTTAACTACCGGAAATGAAC 2598
QY 2530 TTCAAGATGATTAAGCTGAGCTGAGAGTGTGCTGAGTGTCAATGAGACCAAAAT 2589
DB 2599 TTCAAGATGATTAAGCTGAGCTGAGAGTGTGCTGAGTGTCAATGAGACCAAAAT 2598
QY 2590 AAGAAAGTGCAGACATTAACCAAGTGAATGATGATGATGATGATGATGATG 2649
DB 2599 AAGAAAGTGCAGACATTAACCAAGTGAATGATGATGATGATGATGATGATG 2598
QY 2550 TTTCCCTGTAGTCAATGATGAGCAAGAGCAAGCTTTTGTGTGACTGAGCTCGAAGAT 2709
DB 2599 TTTCCCTGTAGTCAATGATGAGCAAGAGCAAGCTTTTGTGTGACTGAGCTCGAAGAT 2603
QY 2710 CTTGCGCTTCCCTGCTGCTACTACAGAGACGTGCAAGATGAGGCGCGGCGGTGAGAA 2769
DB 2604 CTTGCGCTTCCCTGCTGCTACTACAGAGACGTGCAAGATGAGGCGCGGCGGTGAGAA 2663
QY 2770 GCTGCTGGGAGGCTGCTGAGTGTATGAGTGTATGAGTGTATGAGTGTATGAGT 2829
DB 2664 GCTGCTGGGAGGCTGCTGAGTGTATGAGTGTATGAGTGTATGAGTGTATGAGT 2723
QY 2830 CTAATTTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2889
DB 2724 CTAATTTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2783
QY 2890 CAGAGTCAACCCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 2949
DB 2784 CAGAGTCAACCCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 2843
QY 2950 CTTCAATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 3009

```

Db      2844 CCTACATCACTGTAACCTACCTTCTCTCTCAGTGGGAGCAGAGCTCTGCCCCCTT 2903
Oy      3010 GCAGGGAGAGCCCGGTCCTCCGTCGAGTGCAGACCTGAGCTGCTTACGATAGC 3069
Db      2904 GCAGGGAGAGCCCGGTCCTCCGTCGAGTGCAGACCTGAGCTGCTTACGATAGC 2963
Oy      3070 CCGGATAGTGTGATGTTCTTCTTACCCTGAACTTTAAACCTGAAGTAGTAGTAGA 3129
Db      2964 CCGGATAGTGTGATGTTCTTCTTACCCTGAACTTTAAACCTGAAGTAGTAGTAGA 3023
Oy      3130 TGGCTTCTTTTACCCTCTGAGTTTATCACTAGAGTAGTGGCTAAGTACCAAAAA 3189
Db      3024 TGGCTTCTTTTACCCTCTGAGTTTATCACTAGAGTAGTGGCTAAGTACCAAAAA 3083
Oy      3190 ACAAACAAAAACAGAAACAAAAAACAAAAAACCTCAAGCTCTTGTACTGCTAGG 3249
Db      3084 ACAAACAAAAACAGAAACAAAAAACAAAAAACCTCAAGCTCTTGTACTGCTAGG 3143
Oy      3250 TTCAATGCTGCAAAATCACTGAGATTTTGTGTTTAACTAACCCTGCTCCACATTTGCTG 3309
Db      3144 TTCAATGCTGCAAAATCACTGAGATTTTGTGTTTAACTAACCCTGCTCCACATTTGCTG 3203
Oy      3310 GAGATGCTATTTGATGATGTGGGCTCAGATGAGCAAGGTCAAGGGGCCAAAAAATTCC 3369
Db      3204 GAGATGCTATTTGATGATGTGGGCTCAGATGAGCAAGGTCAAGGGGCCAAAAAATTCC 3263
Oy      3370 CCCCTCCCCCAGAGATTTGAAAGATGATTTAAGCTTTAAGCTTCTCTGAGCTT 3429
Db      3264 CCCCTCCCCCAGAGATTTGAAAGATGATTTAAGCTTTAAGCTTCTCTGAGCTT 3323
Oy      3430 CCCCTTGTGTTGTAACAAGGCTGAAGTCTGTTGCTTGTAGCATTTCCAGAGATGAT 3489
Db      3324 CCCCTTGTGTTGTAACAAGGCTGAAGTCTGTTGCTTGTAGCATTTCCAGAGATGAT 3383
Oy      3490 GATGTACAGAGGATGATACATACCACTTTAAGGCTTTTCCCTGGCAGGGGCCCATGTTGG 3549
Db      3384 GATGTACAGAGGATGATACATACCACTTTAAGGCTTTTCCCTGGCAGGGGCCCATGTTGG 3443
Oy      3550 CTAGGCTCAGAGAGCTGAGATGAAATGTTGGAGCTCAGAGAAAGGTGGGTGGAGTGGC 3609
Db      3444 CTAGGCTCAGAGAGCTGAGATGAAATGTTGGAGCTCAGAGAAAGGTGGGTGGAGTGGC 3503
Oy      3610 CCTCTTCCAGGTGTGAGGGATACGAAGAGGAGGAGCTTGGGAAATCCATTCGCCACTGCC 3669
Db      3504 CCTCTTCCAGGTGTGAGGGATACGAAGAGGAGGAGCTTGGGAAATCCATTCGCCACTGCC 3563
Oy      3670 TCTTGCCAATGAGGGGCCAGTCCCAACAGCTCAGGTCCCCAGAACCCCTAGTTCCT 3729
Db      3564 TCTTGCCAATGAGGGGCCAGTCCCAACAGCTCAGGTCCCCAGAACCCCTAGTTCCT 3623
Oy      3730 CATGAGAGCTAGAGCAGCAAGACACATGTCCTTATCTGAGAGAGTGTGGGAACT 3789
Db      3624 CATGAGAGCTAGAGCAGCAAGACACATGTCCTTATCTGAGAGAGTGTGGGAACT 3683
Oy      3790 ACAGTGAAGAACTCTGAGAGTGTAAAGCTTTTACCCACAGATAGTATGTTTAA 3849
Db      3684 ACAGTGAAGAACTCTGAGAGTGTAAAGCTTTTACCCACAGATAGTATGTTTAA 3743
Oy      3850 AGGGGTGCTTTTAAAGGAGCATCACTGAGATGAAGAAAGTGCATTTTCAGAAATGCCAT 3909
Db      3744 AGGGGTGCTTTTAAAGGAGCATCACTGAGATGAAGAAAGTGCATTTTCAGAAATGCCAT 3803
Oy      3910 CGTATATGCTTTTAAAGCCTTTTAACTAATTTACAGTGTCTATTTTATAGAAAGCAGAA 3969
Db      3804 CGTATATGCTTTTAAAGCCTTTTAACTAATTTACAGTGTCTATTTTATAGAAAGCAGAA 3863
Oy      3970 CACTTCTTTTATGACTCTCAGACTTCAATTTTACATTTTATACATTTTGTGTAAGTGC 4029
Db      3864 CACTTCTTTTATGACTCTCAGACTTCAATTTTACATTTTATACATTTTGTGTAAGTGC 3923
Oy      4030 AAGGTGTGGCTTTTGTAACTTCAACAGGTGTGGGAGAGACTGCCCTTGTTCACAGTTT 4089

```

```

Db      3924 AAGGTGTGGCTTTTGTAACTTCAACAGGTGTGGGAGAGACTGCCCTTGTTCACAGTTT 3983
Oy      4090 GTCTCAGCTGTTTCTAATTTTGTAGTGCAGAAAGATGACAGATGCCAGATTTACCTTTC 4149
Db      3984 GTCTCAGCTGTTTCTAATTTTGTAGTGCAGAAAGATGACAGATGCCAGATTTACCTTTC 4043
Oy      4150 TGGTGTAAAGTTGTATTTCTCMAAAAAAAAAAAAAAAAAAAAAA 4195
Db      4044 TGGTGTAAAGTTGTATTTCTCTAAAAAAAAAAAAAAAAAAAAAA 4089

RESULT 11
AF151975      4034 bp mRNA linear ROD 08-JUN-2000
LOCUS        Mus musculus DNA cytosine-specific methyltransferase isoform 7
DEFINITION  AF151975
ACCESSION   AF151975
VERSION     AF151975.1 GI:8347134
KEYWORDS    Mus musculus.
SOURCE      Mus musculus.
ORGANISM    Mus musculus.
REFERENCE   1 (bases 1 to 4034)
AUTHORS     Yin,B., Chen,Y.T., Zhu,M., Luo,Y.J., Zhu,N., Xu,S.C., Wu,G.Y. and
            Shen,Y.
TITLE       Cloning of full-length Dnmt3b cDNA and its alternative splicing
            isoforms in mouse embryonic tissue
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 4034)
AUTHORS     Yin,B., Chen,Y.T., Zhu,M., Luo,Y.J., Zhu,N., Xu,S.C., Wu,G.Y. and
            Shen,Y.
TITLE       Direct Submission
JOURNAL     Submitted (17-MAY-1999) Department of Biochemistry and Molecular
            Biology, Institution of Basic Medical Sciences, Chinese Academy of
            Medical Sciences, Dong Dan San Tiao 5, Beijing City 100005,
            P.R.China

FEATURES
source      1..4034
            /organism="Mus musculus"
            /strain="KM"
            /db_xref="taxon:10090"
            /dev_stage="8-9 day old embryo"
            1..4034
            /gene="Dnmt3b"
            /gene="Dnmt3b"
            /feature="Dnmt3b"
            /function="methylates cytosine in DNA"
            /note="alternatively spliced"
            /codon_start=1
            /product="DNA cytosine-specific methyltransferase isoform
            7"
            /protein_id="AAF74521.1"
            /db_xref="GI:8347135"
            /translation="MKGDSRLHNEEGAGSGYEELIIVGNFSDQSDPTKDAAPSPVLE
            AICEPVCTPETRGRSSRSLSKREVSLNLTGDMTGCDGDEVDGNSDILMPK
            LTRTKQTRTSSESAPVTRHSNQTSSLSLRASRAPIRTGSGHDEQVYVEPAR
            SRRRRASSASATPMSSPASVYPMREYFKTSYSDVLSQGDGDEGDTQVDAESTY
            GDSTIEQDDKEFGIDGLWVKIKGSMVPAVAVSMKTKSFQAMPGRWVQVDFGKF
            SEISADKLVALGLFSQHEFNLTAFNKLVSYYRVAAMHTLEKAVRAGKTFSSSPGSLD
            OLKPLMEAHAGCFPTGIEGLKPKKQVYRVKSKVRSDSHNLEPPRRRBNKSRRTT
            DSAASESPRKLRTNYSYGKRDGEDESRRMASEVYNNKCNLEDCISGKRNPVS
            FHLPEGLDQSCDRFLELYMDDEGYSYCVCCGRELILCSMTSCRCRQVRC
            LEVYAGCTAARDATQDQWMSYKMLPQRCQHVLRKKDMNRLODFTPTDPLDEPFP
            PKLPAIPAKRRPRTIVLSLDGATGTLVLEKGIKYEKTIASEVCAESTIAVTVKH
            EGQIKYVDVVRKIRTKNLEMGPDVLVIGSPCDLSLVNVPARGGLVETGRLEFEY
            HILMYTRKEDNDPFFWMEFENVAMKVDKDISRFLACNPVIDAIKYSAAHRRARY
            FWGNLPGNRIIFGPAPHYTDVSNMGRGAROKLGRSMSPVIRHLFAPLQDYFACE"

BASE COUNT  991 a 1042 c 1097 g 904 t

Query Match      84.7% Score 3554.8 DB 10: Length 4034:
Best Local Similarity 92.7% Pred. No. 0;

```


OY	2170	CCGGAATAATCCAAAGAAATATTTGAAGATGGGGCCGTTGGACATTGGATGGG	2229
Db	2198	CCGGAATAATCCAAAGAAATATTTGAAGATGGGGCCGTTGGACATTGGATGGG	2257
OY	2230	AAGCCATGCAATATCTCTCTAAAGTCAATCTCGCCCGAAAGTTTATATGAGGCAC	2289
Db	2258	AAGCCATGCAATATCTCTCTAAAGTCAATCTCGCCCGAAAGTTTATATGAGGCAC	2317
OY	2290	AGGAAGGCTCTTCTCGAGTTTACACCTGCTGAATTAACCCGCCCAAGAGGCGCA	2349
Db	2318	AGGAAGGCTCTTCTCGAGTTTACACCTGCTGAATTAACCCGCCCAAGAGGCGCA	2377
OY	2350	CAACCGTCCATCTTCTTGAGATGTCGAGAATGTGTGGCCATGAAAGTGAACAA	2409
Db	2378	CAACCGTCCATCTTCTTGAGATGTCGAGAATGTGTGGCCATGAAAGTGAACAA	2437
Db	2410	AGACATCTCAAGATTCTGGCATTAACCCAGTATGATGCATGCATCAAGATGTCTGC	2469
OY	2438	AGACATCTCAAGATTCTGGCATTAACCCAGTATGATGCATGCATCAAGATGTCTGC	2497
Db	2498	TGCTACAGAGGCGCGGACTCTTGGGGTAACTACCGGAATGAACAGGCGCGTATGCG	2529
OY	2530	TTCAAAGATGATTAAGCTCGAGCTGCAGAGACTGCTGAGTTCAGTAGAGACGAAAGTT	2589
Db	2544	-----	2543
OY	2590	AAAGAAAGTGACAGACATTAACCAACAGTGAATCCATCAGACAGGCGAAACCAAGCT	2649
Db	2544	-----	2543
OY	2650	TTTCCCTGATCATGAATGGCAAGACAGACTTTTGTGTGACATGAGCTCGAAAGAT	2709
Db	2544	-----	2543
OY	2710	CTTGGGCTCCCTGCTCACTACAGAGAGTGTCCAAATGGGCGCGCGCCCTCAGAA	2769
Db	2549	CTTGGGCTCCCTGCTCACTACAGAGAGTGTCCAAATGGGCGCGCGCCCTCAGAA	2608
OY	2770	GCTGCTGGGCGAGTCTGAGTGTACCGGATACAGACACTGTTTGGCCCTTGAAGGA	2829
Db	2609	GCTGCTGGGCGAGTCTGAGTGTACCGGATACAGACACTGTTTGGCCCTTGAAGGA	2668
Db	2830	CTACTTGGCTGTAATAGTCTACCCAGGACTGGGAGCTCTGGTCTAGAGCCAGTGC	2889
Db	2669	CTACTTGGCTGTAATAGTCTACCCAGGACTGGGAGCTCTGGTCTAGAGCCAGTGC	2728
OY	2880	CAGAGTACCCCTCCCTGAGAGCACCTACCTGTCCCTTTTATAGTCACTGTGTGGG	2949
Db	2729	CAGAGTACCCCTCCCTGAGAGCACCTACCTGTCCCTTTTATAGTCACTGTGTGGG	2788
OY	2950	CCTACATCACTGTACCTGACCTTCTCTCTCAGTGGGAGCAGAGCTCTGGCCCTT	3009
Db	2789	CCTACATCACTGTACCTGACCTTCTCTCTCAGTGGGAGCAGAGCTCTGGCCCTT	2848
OY	3010	GCAGGGGAGCCCGGTGCTCCCTCCCTGCTGACAGCTCAGACCTGAGCTTAAAGTAC	3069
Db	2849	GCAGGGGAGCCCGGTGCTCCCTCCCTGCTGACAGCTCAGACCTGAGCTTAAAGTAC	2908
OY	3070	CCGGCATGTGTCTATGTCTCTTACCTGAACTTTAAACTTGAAGTAGTAAAGA	3129
Db	2909	CCGGCATGTGTCTATGTCTCTTACCTGAACTTTAAACTTGAAGTAGTAAAGA	2968
OY	3130	TGGCTTTCTTACCTCTCAGATTATCACTCAAGAGTGTGGTAAAGTAAACAAAAA	3189
Db	2969	TGGCTTTCTTACCTCTCAGATTATCACTCAAGAGTGTGGTAAAGTAAACAAAAA	3028
OY	3190	ACAAACAAAAACAGAAACAAAAAACAAAAAACCTCAACAGCTCTTTAGTACTAGG	3249
Db	3029	ACAAACAAAAACAGAAACAAAAAACAAAAAACCTCAACAGCTCTTTAGTACTAGG	3088
OY	3250	TTCAATGCTGCAAAATCACTGAGATTTTGTTTTAAAGTAAACCGGTGCTCCACATTTGCTG	3309

Db	3089	TTCAATGCTGCAAAATCACTTGAATTTTGTATTAAGTAACCCGTGCTCCACATTTGCTG	3148
QY	3310	GAGATGCTATTTGTAATGTGGGCTCGAGATGAGCAAGGTCAAAGGGGCCAAAAAAATTC	3369
Db	3149	GAGATGCTATTTGTAATGTGGGCTCGAGATGAGCAAGGTCAAAGGGGCCAAAAAAATTC	3208
QY	3370	CCCTCTCCCGCCAGAGATATTTGAAGATGATGTTTATGTTTAAAGTCTCTCGCACCTT	3429
Db	3209	CCCTCTCCCGCCAGAGATATTTGAAGATGATGTTTATGTTTAAAGTCTCTCGCACCTT	3268
QY	3430	CCCTCTGCTTGTGTACAAGGGCTGAAGCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	3489
Db	3269	CCCTCTGCTTGTGTACAAGGGCTGAAGCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	3328
QY	3490	GATGTCCAGCAGGATGACATCACCACTTTAGGCTTTTCCCTGGCAGAGGGCCCATGTGG	3549
Db	3329	GATGTCCAGCAGGATGACATCACCACTTTAGGCTTTTCCCTGGCAGAGGGCCCATGTGG	3388
QY	3550	CTAGTCCTCAGCAAGAGACTGAGTGAATGTTTGGAGCTCAGAAAGGGTGGGTGAGTGGC	3609
Db	3389	CTAGTCCTCAGCAAGAGACTGAGTGAATGTTTGGAGCTCAGAAAGGGTGGGTGAGTGGC	3448
QY	3610	CCTCTTCCAGGTGTGAGGGATACCAAGAGAGAACTTATGGGAATCCATTCCCATCTCC	3669
Db	3449	CCTCTTCCAGGTGTGAGGGATACCAAGAGAGAACTTATGGGAATCCATTCCCATCTCC	3508
QY	3670	TCTTGGCCAAATGAGAGGGGCCAGTCCCAACAGCTCAGGTGCCCAAGAACCCCTAGTTCCT	3729
Db	3509	TCTTGGCCAAATGAGAGGGGCCAGTCCCAACAGCTCAGGTGCCCAAGAACCCCTAGTTCCT	3568
QY	3730	CATGAGAAGCTAGGACCAAGAACCAATGTTCCCTTATCTGAGCAGTGTGTTGGGGAAT	3789
Db	3569	CATGAGAAGCTAGGACCAAGAACCAATGTTCCCTTATCTGAGCAGTGTGTTGGGGAAT	3628
QY	3790	ACAGTGAAGAACTTCTGAGATGTTAAAGCTTTTACCACAGATAGATTGTGTGTTTA	3849
Db	3629	ACAGTGAAGAACTTCTGAGATGTTAAAGCTTTTACCACAGATAGATTGTGTGTTTA	3688
QY	3850	AGGGGTGCTTTTATGAGGGGCAATCACTGAGATAAAGAAAGCTGCATTTACAAATGCCAT	3909
Db	3689	AGGGGTGCTTTTATGAGGGGCAATCACTGAGATAAAGAAAGCTGCATTTACAAATGCCAT	3748
QY	3910	CGTAATGCTTTTAAACACCTTTTACCTAATTTACAGGTGCTATTTATAGAAGCAGACAA	3969
Db	3749	CGTAATGCTTTTAAACACCTTTTACCTAATTTACAGGTGCTATTTATAGAAGCAGACAA	3808
QY	3970	CACCTCTTTTATGACTCTGAGACTTCTATTTTCATGTACCATTTTTTTTGTAACTGCG	4029
Db	3809	CACCTCTTTTATGACTCTGAGACTTCTATTTTCATGTACCATTTTTTTTGTAACTGCG	3868
QY	4030	AAGGTGTGGGCTTTTGTAACTTCAACAGGTGTGGGAGAGAGCTGCTTTTCAACAGTTT	4089
Db	3869	AAGGTGTGGGCTTTTGTAACTTCAACAGGTGTGGGAGAGAGCTGCTTTTCAACAGTTT	3928
QY	4090	GTCCTCCACTGGTTTCTAATTTTATAGGTGCAAAAGATGACAGATGCCACAGTTTACCTTC	4149
Db	3929	GTCCTCCACTGGTTTCTAATTTTATAGGTGCAAAAGATGACAGATGCCACAGTTTACCTTC	3988
QY	4150	TGCTTGAATTAAGTTGTATTTCTCTAATAAAAAAAAAAAAAAAAAAAAAA 4195	
Db	3989	TGCTTGAATTAAGTTGTATTTCTCTAATAAAAAAAAAAAAAAAAAAAAAA 4034	
RESUL.T 12			
AF151976			
LOCUS			
DEFINITION Mus musculus DNA cytosine-specific methyltransferase isoform 8			
ACCESSION AF151976			
VERSION AF151976.1 GI:8347136			
KEYWORDS Mus musculus.			
SOURCE			

ORGANISM	Mus musculus
REFERENCE	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	Yin, B., Chen, Y.T., Zhu, M., Luo, Y.J., Zhu, N., Xu, S.C., Wu, G.Y. and Shen, Y.
TITLE	Cloning of full-length Dnmt3b cDNA and its alternative splicing
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 3974)
AUTHORS	Yin, B., Chen, Y.T., Zhu, M., Luo, Y.J., Zhu, N., Xu, S.C., Wu, G.Y. and Shen, Y.
TITLE	Direct Submission
JOURNAL	Submitted (17-MAY-1999) Department of Biochemistry and Molecular Biology, Institution of Basic Medical Sciences, Chinese Academy of Medical Sciences, Dong Dan San Tiao 5, Beijing City 100005, P.R.China
FEATURES	Location/Qualifiers
source	1..3974
gene	/organism="Mus musculus"
CDS	/strain="KM"
	/db_xref="taxon:10090"
	/dev_stage="8-9 day old embryo"
	1..3974
	/gene="Dnmt3b"
	297..2627
	/gene="Dnmt3b"
	/function="methylates cytosine in DNA"
	/note="alternatively spliced"
	/codon_start=1
	/product="DNA cytosine-specific methyltransferase isoform 8"
	/protein_id="AAF4522.1"
	/db_xref="GI:8347137"
	/translation="MKGDSRHLNEEGAGYEETITVNGNTPSDSDTKDAPAPVLE ALCTPCTPTGRSSRSRLSKREVSLILNTYQMTGDGRDDEVDDGSDILMPK LRETQKTRTESPAVRTRHNSGTSLSLROKASPRITRGOGHHVOEYEPFAPTR SRRRASSASTPMSPVASVDPEMEVTPKVSSTPSVDLSODGDEGDTTOVDAEISY GSTYDODKEFGIDLVMGKIKFPMAMVAVSWKATSKQAPGMRKQWOMQDGEK SEISADKLVALGFSQHNPLATNKLIVSRKAMHTLTKAVRAKGTFFSSPQESIED QLKPLMAWAGFPKPTGIEGLKPNKQPKNSRRRTYDASAPSPKTKTNSYGG KRDGEDSRERMAEVTNNKMLNEDRCLSCCKKPVGFHPLFPGGLCOSQDRPDEL FYMYDEYQSYCTVCCCEGRELILCSNTSCRCFVECELEVLVAGAEPAEKIQEWS CYMCIPORCHGVLRKRDMMRLQDFETDLPDELEFEPKLYPAIPAKRRPTRIVSL PGIATGVLVLEKELIKERYIASEVCAESIAGVFKHEGQIKYVNDVRKIKTKNLE MCPDLVIGSPNDLSNVPARKGLYEGRGLFEFPHLNTYRPREGNRPFFYMF EVNVAMKYNDDKDISRLACNPYMDIAIVSAHARFVGMNPLPMNRIFGFPAYHTD VENMGRAROKLGRSWSVPIRHLPAPEKDYFACE"
BASE COUNT	972 a 1032 c 1078 g 892 t
ORIGIN	
Query Match	81.6%; Score 3424.8; DB 10; Length 3974;
at Local Similarity	93.5%; Pred. No. 0;
Matches 3689; Conservative	0; Mismatches 7; Indels 249; Gaps 2;
QY	251 GCCTTTTACAGAAACATGAAGGAGACAGACATCTGAATGAGAGAGAGGCGCC 310
DB	551 GGCCTGATATTCTAAATGCAAAAGCTCACCCTGAGACCAAGACACAGAGCGCTCT 610
QY	279 GCCTACAGACAGAAACATGAAGGAGACAGACATCTGAATGAGAGAGGCGCC 338
DB	579 GGCCTGATATTCTAAATGCAAAAGCTCACCCTGAGACCAAGACACAGAGCGCTCT 638
QY	311 AGCGGATAGAGATGATATGCTTAATGGAACCTCAGAGACATCTCAGACAGC 370
DB	611 GAAAGCGCGGCTGCCGAAACCCGACATAGATGGGACCTCCAGCTTGAGAGCAAA 670
QY	339 AGCGGATAGAGATGATATGCTTAATGGAACCTCAGAGACATCTCAGACAGC 398
DB	639 GAAAGCGCGGCTGCCGAAACCCGACATAGATGGGACCTCCAGCTTGAGAGCAAA 698
QY	371 AAGGATGCTCCCTCAGCCCAAGCTTCTGAGAGCAATCTGACAGACAGCAGTGCACACA 430
DB	671 GCTTCCCGGACCAATCCCGAGGCTCGGACAGGCGCGCACATGTGACAGAGTACCTGTG 730
QY	399 AAGGATGCTCCCTCAGCCCAAGCTTCTGAGAGCAATCTGACAGACAGCAGTGCACACA 458
DB	699 GCTTCCCGGACCAATCCCGAGGCTCGGACAGGCGCGCACATGTGACAGAGTACCTGTG 758
QY	431 GAGACAGAGGCGGAGGTCAAGCTCCGGCTCTTAAGAGGAGGTCTCAGACCTCTG 490
DB	731 GAGTTTCGGCTACAGAGTCTCGGACGCTGACGCTGCTTCAGCAAGACAGCCATG 790
QY	459 GAGACAGAGGCGGAGGTCAAGCTCCGGCTCTTAAGAGGAGGTCTCAGACCTCTG 518
DB	759 GAGTTTCGGCTACAGAGTCTCGGACGCTGACGCTGCTTCAGCAAGACAGCCATG 818
QY	491 AATTACCCGAGACATGACAGAGATGAGAGAGAGATGATGAGATGAGATGAGAT 550
DB	791 TCATCCCTCGGACGCTGCTGATCATGAAAGTGACACCTAAGAGGCGTACACCCCA 850
QY	1539 AATCTGGAAGACCGCTGTGTCTGTGTAAGAAAGAACCTGTGTCTTCCACCCCTTC 1630
DB	819 TCATCCCTCGGACGCTGCTGATCATGAAAGTGACACCTAAGAGGCGTACACCCCA 878
QY	1571 AATCTGGAAGACCGCTGTGTCTGTGTAAGAAAGAACCTGTGTCTTCCACCCCTTC 1630
DB	851 TCAGTTGACTTGAGGACAGATGAGATGAGAGGAGATGAGATGAGATGAGATGAG 910
QY	1598 AATCTGGAAGACCGCTGTGTCTGTGTAAGAAAGAACCTGTGTCTTCCACCCCTTC 1630
DB	879 TCAGTTGACTTGAGGACAGATGAGATGAGAGGAGATGAGATGAGATGAGATGAG 938
QY	911 GAGAGAGAGATGAGAGAGACAGACAGATGATGAGATGAGATGAGATGAGATGAG 970
DB	939 GAGAGATGATGAGAGAGACAGACAGATGATGAGATGAGATGAGATGAGATGAG 998
QY	971 CTCTGTTGGGAAAGATCAAGGCTTCTCTGTTGGGCTCGCATGTTGTTGTTGTTG 1030
DB	999 CTCTGTTGGGAAAGATCAAGGCTTCTCTGTTGGGCTCGCATGTTGTTGTTGTTG 1058
QY	1031 GGCACCTCCAAAGCAGACAGGCTTCCCGGATGCGGTGGTACATGTTGGTGGATG 1090
DB	1059 GGCACCTCCAAAGCAGACAGGCTTCCCGGATGCGGTGGTACATGTTGGTGGATG 1118
QY	1091 AAGTTTTCGAGATCTGCTGCTGACAAATGAGGCTCTGGGGCTGTTACAGCGACACT 1150
DB	1119 AAGTTTTCGAGATCTGCTGCTGACAAATGAGGCTCTGGGGCTGTTACAGCGACACT 1178
QY	1151 AATCTGAGTACCTTCAATAGAGCTGTTTCTTATAGAGGCGCATGATACACACTGTGAG 1210
DB	1179 AATCTGAGTACCTTCAATAGAGCTGTTTCTTATAGAGGCGCATGATACACACTGTGAG 1238
QY	1211 AAGCCAGGTTGCGAGTGGCAGAGACCTTCTCAGAGCTCTGAGAGTCACTGAGAGAC 1270
DB	1239 AAGCCAGGTTGCGAGTGGCAGAGACCTTCTCAGAGCTCTGAGAGTCACTGAGAGAC 1298
QY	1271 CAGCTGAAGCCATGCTGAGTGGGCGCCACAGGTTGCTTCAAGCCCTACTGGATCGAGG 1330
DB	1299 CAGCTGAAGCCATGCTGAGTGGGCGCCACAGGTTGCTTCAAGCCCTACTGGATCGAGG 1358
QY	1331 CTCAAAACCCAAAGAAAGCAACAGTGGTTAATAGTGAAGTGGCGTCTGACAGAGT 1390
DB	1359 CTCAAAACCCAAAGAAAGCAACAGTGGTTAATAGTGAAGTGGCGTCTGACAGAGT 1381
QY	1391 AAGAACTTGAACCCAGAGAGACCGAAGCAAAAGTGCAGAAAGCAGCAACCAATGACTCT 1450
DB	1382 AAGAACTTGAACCCAGAGAGACCGAAGCAAAAGTGCAGAAAGCAGCAACCAATGACTCT 1418
QY	1451 GCTGCTTCTGAGTCCCGCCCAACCAAGCGCTTCAAGACAAATAGCTTATGCGGGAAGGAC 1510
DB	1419 GCTGCTTCTGAGTCCCGCCCAACCAAGCGCTTCAAGACAAATAGCTTATGCGGGAAGGAC 1478
QY	1511 CGAGGAGAGATGAGAGAGCGCAGAAAGATGAGCTTCTGAAAGTACCAACCAAGAGG 1570
DB	1479 CGAGGAGAGATGAGAGAGCGCAGAAAGATGAGCTTCTGAAAGTACCAACCAAGAGG 1538
QY	1571 AATCTGGAAGACCGCTGTGTCTGTGTAAGAAAGAACCTGTGTCTTCCACCCCTTC 1630
DB	1539 AATCTGGAAGACCGCTGTGTCTGTGTAAGAAAGAACCTGTGTCTTCCACCCCTTC 1598

QY 1631 TTGAGGGTGGGCTCTGTGACAGAGTTGCCGGGATCGTTCCTAGAGCTCTTCTACATGAT 1690
 |||||
 Db 1599 TTGAGGGTGGGCTCTGTGACAGAGTTGCCGGGATCGTTCCTAGAGCTCTTCTACATGAT 1658
 QY 1691 GATGAGAGCGGCTATCAGTCTCTACTGCACCGCTGTCTGTGAGGGCCGTGAACCTGCTG 1750
 |||||
 Db 1659 GATGAGAGCGGCTATCAGTCTCTACTGCACCGCTGTCTGTGAGGGCCGTGAACCTGCTG 1718
 QY 1751 TGCAGTAAACAACTCTGTGAGATGCTGTCTGTGAGAGTGTCTGTGAGAGTGTCTGTGAGG 1810
 |||||
 Db 1719 TGCAGTAAACAACTCTGTGAGATGCTGTCTGTGAGAGTGTCTGTGAGAGTGTCTGTGAGG 1778
 QY 1811 GCAGCAGAGCTGAGATGCAAGCTGCAAGAGACCTGAGAGCTGCTATATGTGCTCCCT 1870
 |||||
 Db 1779 GCAGCAGAGCTGAGATGCAAGCTGCAAGAGACCTGAGAGCTGCTATATGTGCTCCCT 1838
 QY 1871 CAGCGCTGCCATGGGGTCTCCGACGCAAGAAAGATTGGAACATGCGCTGCAAGACTTC 1930
 |||||
 Db 1839 CAGCGCTGCCATGGGGTCTCCGACGCAAGAAAGATTGGAACATGCGCTGCAAGACTTC 1898
 QY 1931 TTCAGTCTGATCTCTGACCTGGAAGAAATTGAGCACCCCAAGTTGTACCCGCAATTCCT 1990
 |||||
 Db 1899 TTCAGTCTGATCTCTGACCTGGAAGAAATTGAGCACCCCAAGTTGTACCCGCAATTCCT 1958
 QY 1991 GCAGCAAAAGAGAGCCCATTTAGAGTCTCTGTCTGTGATGGAATTGCAACGGGGTAC 2050
 |||||
 Db 1959 GCAGCAAAAGAGAGCCCATTTAGAGTCTCTGTCTGTGATGGAATTGCAACGGGGTAC 2018
 QY 2051 TTGCTGCTCAAGAGATGGGTATTAAAGTGGAAAAAGTACATTCCTCCGAAAGTCTGCA 2110
 |||||
 Db 2019 TTGCTGCTCAAGAGATGGGTATTAAAGTGGAAAAAGTACATTCCTCCGAAAGTCTGCA 2078
 QY 2111 GAGTCACATGCTGTGGGAACTGTAAGATGAAGCCAGATCAATAATGTCAATGAGCTC 2170
 |||||
 Db 2079 GAGTCACATGCTGTGGGAACTGTAAGATGAAGCCAGATCAATAATGTCAATGAGCTC 2138
 QY 2171 CGGAAAAATCACCAAGAAAAATATTGAAGAGTGGGGCCGTTCGACTTGTGATTTGTTGGA 2230
 |||||
 Db 2139 CGGAAAAATCACCAAGAAAAATATTGAAGAGTGGGGCCGTTCGACTTGTGATTTGTTGGA 2198
 QY 2231 AGCCCATGCAATGATCTCTTACGTCATCTCCGCGCAAAAGTTTATATGAGGGCACA 2290
 |||||
 Db 2199 AGCCCATGCAATGATCTCTTACGTCATCTCCGCGCAAAAGTTTATATGAGGGCACA 2258
 QY 2291 GGAAGGCTCTCTGAGTTTACACTGCTGTAATTATACCGGCCCAAGAGGGCGAC 2350
 |||||
 Db 2259 GGAAGGCTCTCTGAGTTTACACTGCTGTAATTATACCGGCCCAAGAGGGCGAC 2318
 QY 2351 AACGTCATCTCTGATGTTGAGAAATGTTGTGGCATGAAGTAAGTACAGAGAA 2410
 |||||
 Db 2319 AACGTCATCTCTGATGTTGAGAAATGTTGTGGCATGAAGTAAGTACAGAGAA 2378
 QY 2411 GACATCTCAGATCTCTGCGCATGTAAACCAGTATGATGATGCCATCAAGGTGTCTCT 2470
 |||||
 Db 2379 GACATCTCAGATCTCTGCGCATGTAAACCAGTATGATGATGCCATCAAGGTGTCTCT 2438
 QY 2471 GCTACAGGGGCCGGTACTCTGGGGTAACTTACCCGGAATGAACAGAGCCCGTGTGCT 2530
 |||||
 Db 2439 GCTACAGGGGCCGGTACTCTGGGGTAACTTACCCGGAATGAAC----- 2483
 QY 2531 TCAGAGATGATTAAGCTCGAGCTCGAGACTGCTGTGAGTTCAGTAGAGCAGCAAAAGTTA 2590
 |||||
 Db 2484 ----- 2483
 QY 2591 AAGAAATGCGAGACATTAACCAACAGTGCATCTCATCAGACAGGGCAAAACCAAGCTT 2650
 |||||
 Db 2484 ----- 2483
 QY 2651 TTCCTGTATGATGAATGGCAAGAGACGTTTTTGTGTGCTGACCTGAGCTCGAAGAGTTC 2710
 |||||
 Db 2484 -----AGGATC 2489

QY 2711 TTGCGCTTCCCTGCTCACTACACGAGGTGTCCAAATGGGCCCGCGGCCCTGACAGAG 2770
 |||||
 Db 2490 TTGCGCTTCCCTGCTCACTACACGAGGTGTCCAAATGGGCCCGCGGCCCTGACAGAG 2549
 QY 2771 CTGCTGGGAGGCTCTGAGATGTACCGGTCAATCAGACACTGTTTCCCTTGAAGAC 2830
 |||||
 Db 2550 CTGCTGGGAGGCTCTGAGATGTACCGGTCAATCAGACACTGTTTCCCTTGAAGAC 2609
 QY 2831 TACTTGGCTGTGAATGTTCTTACCCAGAGAGTGGGGAGCTGTGCTGAGAGCCAGTCCG 2890
 |||||
 Db 2610 TACTTGGCTGTGAATGTTCTTACCCAGAGAGTGGGGAGCTGTGCTGAGAGCCAGTCCG 2669
 QY 2891 AGAGTCAACCCCTCCCTGAAGGACCTCACTGCTCCCTTTTACCTCACTGTGTGGGC 2950
 |||||
 Db 2670 AGAGTCAACCCCTCCCTGAAGGACCTCACTGCTCCCTTTTACCTCACTGTGTGGGC 2729
 QY 2951 CTCACTACCTGTATCTGAGCTTCTCTGCTGCTGAGTGGGAGAGAGCTCTGCTGCTTG 3010
 |||||
 Db 2730 CTCACTACCTGTATCTGAGCTTCTCTGCTGCTGAGTGGGAGAGAGCTCTGCTGCTTG 2789
 QY 3011 CAGGGAGAGCCCGGTGCTCCCTCCGTGAGACACTGAGACCTGGCTGTAGAGTACC 3070
 |||||
 Db 2790 CAGGGAGAGCCCGGTGCTCCCTCCGTGAGACACTGAGACCTGGCTGTAGAGTACC 2849
 QY 3071 CGGATGAGTGTCTATGTTCTTACCTGAAACTTTAAACTGAAGTAGTAGTAAAGT 3130
 |||||
 Db 2850 CGGATGAGTGTCTATGTTCTTACCTGAAACTTTAAACTGAAGTAGTAGTAAAGT 2909
 QY 3131 GGCCTTTCTTTACCTCTCTGAGTTTATCACTCAGAAAGTATGGCTAAGATACCAAAAAA 3190
 |||||
 Db 2910 GGCCTTTCTTTACCTCTCTGAGTTTATCACTCAGAAAGTATGGCTAAGATACCAAAAAA 2969
 QY 3191 CAACCAAAAAAGAGAAACAAAAAACAACCTCAACAGCTCTCTAGTACTCAGT 3250
 |||||
 Db 2970 CAACCAAAAAAGAGAAACAAAAAACAACCTCAACAGCTCTCTAGTACTCAGT 3029
 QY 3251 TCATGCTGCAAAATCACTGAGATTTGTTTAAAGTAAACCCGTCTCCACATTTGCTGG 3310
 |||||
 Db 3030 TCATGCTGCAAAATCACTGAGATTTGTTTAAAGTAAACCCGTCTCCACATTTGCTGG 3089
 QY 3311 AGGATGCTATTGTCAATGTGGGCTCAGATGAGCAAGGTCAAGGGGCCAAAAAATTC 3370
 |||||
 Db 3090 AGGATGCTATTGTCAATGTGGGCTCAGATGAGCAAGGTCAAGGGGCCAAAAAATTC 3149
 QY 3371 CCTCTCCCCAGGAGTATTTGAAGATGATTTATGTTTAAAGTCTCCGAGACCTTC 3430
 |||||
 Db 3150 CCTCTCCCCAGGAGTATTTGAAGATGATTTATGTTTAAAGTCTCCGAGACCTTC 3209
 QY 3431 CCTCTGCTTGTGTAACAGGGCTGAAGTCTGTGCTTGTAGCATTTCCAGGATGATG 3490
 |||||
 Db 3210 CCTCTGCTTGTGTAACAGGGCTGAAGTCTGTGCTTGTAGCATTTCCAGGATGATG 3269
 QY 3491 ATGTGAGAGGATGATCACTACACACTTTTAGGGCTTTTCCCTGGCAGGGGCCATGTGGC 3550
 |||||
 Db 3270 ATGTGAGAGGATGATCACTACACACTTTTAGGGCTTTTCCCTGGCAGGGGCCATGTGGC 3329
 QY 3551 TAGTCTCAGCAAGCTGAGTATGATTTTGAAGTGTGAGAGTCAAGAGGTTGGGTGAGTGGC 3610
 |||||
 Db 3330 TAGTCTCAGCAAGCTGAGTATGATTTTGAAGTGTGAGAGTCAAGAGGTTGGGTGAGTGGC 3389
 QY 3611 CTCTTCAGGATGTAGAGGATACGAAGAGGAAGTTAGGAATTCATTTCCCACTCCCT 3670
 |||||
 Db 3390 CTCTTCAGGATGTAGAGGATACGAAGAGGAAGTTAGGAATTCATTTCCCACTCCCT 3449
 QY 3671 CTTGCGCAATGAGGGGCCCAATCCCAACACTCAAGTGTCCCAAGAACCCCTAATTCTC 3730
 |||||
 Db 3450 CTTGCGCAATGAGGGGCCCAATCCCAACACTCAAGTGTCCCAAGAACCCCTAATTCTC 3509
 QY 3731 ATGAGAGCTAGGAGCAAGAACATGCTTCCCTTATCTGAGAGTGTGTTGGGAACCTA 3790
 |||||
 Db 3510 ATGAGAGCTAGGAGCAAGAACATGCTTCCCTTATCTGAGAGTGTGTTGGGAACCTA 3569
 QY 3791 CAGTGAACCACTTCTGAGATGTTAAAAAGCTTTTACCCACAGATAGATTGTGTTTAA 3850

Db	3570	CAGTGAAAAACCTCTGTGAGATGTTAAAGCTTTTTTACCCACGATAGATGTGTTTTTAA	36282
OY	3851	GGGGTGTCTTTTTTAAAGGGGCATCACTGAGATTAAGAAACCTGCATTTCCAAAATGGCATC	39110
Db	3630	GGGGGCTTTTTTTTAAAGGGGCATCACTGAGATTAAGAAACCTGCATTTCCAAAATGGCATC	36893
OY	3911	GTAATGCTTTTTTAAACACCTTTTACCTAATTTACAGGTGCTATTTTATAGAAGCAGAAC	39707
Db	3690	GTAATGCTTTTTTAAACACCTTTTACCTAATTTACAGGTGCTATTTTATAGAAGCAGAAC	37489
OY	3971	ACTTCTTTTTATGACTCTCAGACTTCTAATTTTCATGTTTACCATTTTTTTTGTAACTCGCA	40330
Db	3750	ACTTCTTTTTATGACTCTCAGACTTCTAATTTTCATGTTTACCATTTTTTTTGTAACTCGCA	38093
OY	4031	AGCTGTGGCTTTTTGTAACTTCCACAGGTGGGGAGACAGCTGCTGTTTCAACAGATTGG	40990
Db	3810	AGGCTGTGGCTTTTTGTAACTTCCACAGGTGGGGAGACAGCTGCTGTTTCAACAGATTGG	38653
OY	4091	TCTCCACTGCTTTCTAATTTTTTAAAGTCCAAAGATGACAGATGCCCAAGATTACCTTTCT	41500
Db	3870	TCTCCACTGCTTTCTAATTTTTTAAAGTCCAAAGATGACAGATGCCCAAGATTACCTTTCT	39293
OY	4151	GCTTGATTTAAAGTTGTAATTTCTCTAATAAAAAAAAAAAAAAAAAAAAAA	4195
Db	3930	GCTTGATTTAAAGTTGTAATTTCTCTAATAAAAAAAAAAAAAAAAAAAAAA	3974

[illegible][illegible]

QY	1181	ATATGGAAGGCATATGTACACACACTCTGGAGAAACCAACGAGGTTTCAGCTGGGAAAGACTTC	1240
Db	1213	TATCGAAAGGCATATGTACACACTCTCTGGAGAAACCTATAGGGTGGCAGCTGGGAAACCTTC	1272
QY	1241	TCCAGCAGTCTCTGGAGAGTCACTGTGAGAGACCAGCTGAAAGCCATGCTGAGATGGGGCCAC	1300
Db	1273	CCGAGCAGCCCTGGAGACTATTTGGAGAGACCAGCTGAAAGCCCATGTTGGATGGGGCCAC	1332
QY	1301	GGTGGCTTCAAGCTTACTGGGATCGAGGGCTCAAACCCAAACAAGAACACAGTGGTT	1360
Db	1333	GGGGGCTTCAAGCCCACTGGGGATGAGGGCCCTCAACCCAAACAACAGCAACCACTGGTT	1392
QY	1361	AATATAGTCGAGGTGGCCGCTTCAGACAGTGTGGAACCTTAGAACCCAGAGACGGGAGAC	1420
Db	1393	AATATAGTCGAGAGGTGGCTTCGTGCAGGCACTAGGAATTTGATATAGGAATATACGAGAC	1452
QY	1421	AAAATGCGAAGAGCACAACCAATGACTCTGCTGTTCTG--AGTCCCCCACCAG	1477
Db	1453	AAGACTCGAAGACCCACAGCTGAGCAGCTACACCCACTCTGACTCTGCCCCCGACCAAG	1512
QY	1478	CGCCCTCAAGCAAAATAGCTAT--GGGGGAGAGACCGAGGGGAGATGAGAGAGACGCA	1534
Db	1513	CGCCCTCAAGCAAAATGCTATTAACAGGGCAAAACCGAGGGGATGGAAGATCAAGACGCA	1572
QY	1535	GAAAGGATGGCTTGTGAAGTCACCAACACAAAGGGCAATCTGGAAGACCGCTGTTTGC	1594
Db	1573	GAACAAAATGGCTTCAGATGTTGCCAACAAAGAGACGACCTGGAAAGATGGCTGTTGTCT	1632
QY	1595	TGTGGAAAGAAAGAACCCCTGTGCTCTCCACCCCTCTTTAGGGTGGGCTCTGTCAGACT	1654
Db	1633	TGTGGACAGGAAAAACCCCGTCTCTCCACCCCTCTTTAGGGGGGGGCTCTGTACAGCA	1692
QY	1655	TGCCGGATCCCTCTCTAGAGCTCTTACACTGTATGATGAGAGCGGCTATCACTCCAC	1714
Db	1693	TGCCGGATGGCTTCTCTTGAAGCTGTTTACATGATATGATGACATGCTATCACTCTTAC	1752
QY	1715	TGCACCGTGTCTGTGAGGGCCGTGAAGCTGCTGTGTCAGTAAACAAAGCTGTGCAGA	1774
Db	1753	TGCACGTGTCTGTGAGGGCCGAGAGCTGCTGCTTTGCACAAACAGCACTGCTGCCGG	1812
QY	1775	TGCTTCTGTGGAAGTGTCTGAGAGTGTGTGGGGCGAGGACACAGCTGAGAGATGGCCAG	1834
Db	1813	TGTTTCTGTGTGAGAGTGTCTGAGAGTGTGTGGGGCGAGGACACAGGCGCCGAGGCCAAG	1872
QY	1835	CTGCAGGAACCTGGAGCTGCTATATGTGCTCCCTCCAGCGCTGCATGGGGTCCCTCGA	1894
Db	1873	CTTCAGGAGCCCTGGAGCTGTATACATGTGTCTCCCGACAGCGCTGTANTGGGTCCTGGG	1932
QY	1895	CGCAGGAAGAATTTGGAACATGCGGCTGTGAACCTTCTTCACTACTGATCTGACCTGGAA	1954
Db	1933	CGCCGGAAGACTGGAAAGTGCAGCTGAGGCCCTTCTTCAACAGTGAACCGGGCTTTGAA	1992
QY	1955	GAAATTTGAGCAACCAAGTTTGTACCCACCAATTCCTGAGGCCAAAGAGGCCCATTTAGA	2014
Db	1993	TA---TGAAGCCCCCAAGCTGTACCCTGTCCATTTCCGAGGCCGGAAGGCCCATTTGGA	2049
QY	2015	GTCTGTCTGTGTTGATGAGAAATGCAAGGGGTAATTGTGTCTCAAGGAATTTGGGTAAT	2074
Db	2050	GTCTGTCTATTTGTTGATGTGAGCATGCGCGACAGGCTACCTAAGTCTCAAGAAATTTGGCATA	2109
QY	2075	AAAGTGAAGAAATGATCATGTCCTCGAABTGTGTGAGGTGCATTCGCTGTGGGAACTTT	2134
Db	2110	AAGGTAGAAAGTATGCTGCTTCTGAAATGTGTGTGGAGTGCATTTGCTGTGGAAACCTGT	2169
QY	2135	AAGCATGAAGGCACAGATCAAAATATGTCAATAGCTCGGAAATTCACCAAGAAAAATTT	2194
Db	2170	AAGCAGGAGGGAATATCAAAATAGCTGAACACGCGAGAGAACATCAAAAAGAAATTT	2229
QY	2195	GAAAGTGGGGCCGCTTGCAGCTTGGTATTTGCTGGAAGCCCATGTCAATGATCTCTTAAC	2254
Db	2230	GAAAGATGGGGCCATTTTGTACTTGGTATTTGGCGGAAGCCCATGTCAAGATCTCTCAAT	2289

[illegible]

Db 3351 TGTAGCATGTAGTACGACATTTTAAAGGCCAGCATGTTTTCACAGGAG 3410
OY 3381 CAGAGATATTTGACATGATGTTTAAAGTCTTCTGACACCTTCCCTTCTT 3440
Db 3411 CAGAG-AGAAATGTTGATATGCT-ATTACCCGACATTCCTTCTGCTT 3462
OY 3441 GGTACAGGCTGATGCTTCTTCTG-TCCTGATGATTTCCAGATGATGAT 3493
Db 3463 AATACAGGCTGATGCTGACGACCTATTTAGATTTTCCACATGATGAT 3522
OY 3494 TCAGCAGGATGATGATGACAC 3516
Db 3523 TCAGCAGGATGATGATGATGATC 3545

RESULT 14

AF156488 4145 bp mRNA linear PRI 05-SEP-1999
LOCUS Homo sapiens DNA cytosine-5 methyltransferase 3 beta 1 (DNMT3B)
DEFINITION mRNA, complete cds.
ACCESSION AF156488
VERSION AF156488.1 GI:5823167

ORIGIN

Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
Xie, S., Wang, Z., Okano, M., Nogami, M., Li, Y., He, W. W., Okumura, K.
and Li, E.
Cloning, expression and chromosome locations of the human DNMT3
gene family
Gene 236 (1), 87-95 (1999)
99365304
10433969
2 (bases 1 to 4145)
Xie, S., and Li, E.
Direct Submission
Submitted (03-JUN-1999) CVRC, Mass. Gen. Hospital, 149 13th Street,
Charlestown, MA 02129, USA

REFERENCE

JOURNAL
MEDLINE
PUBMED
AUTHORS
TITLE
JOURNAL

FEATURES
source
1. 4145
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="20"
/map="between D20S182 and D20S106; 39.9-50.2 cM"
1. 4145
/gene="DNMT3B"
115..2676
/note="de novo DNA methyltransferase"
/product="DNA cytosine-5 methyltransferase 3 beta 1"
/protein_id="AAD53063.1"
/db_xref="GI:5823167"
/translation="MKGDYRLNGEDDAGREDSTLVNACDSQSSDPTLEAIRP
EIRGRSSRLKREVSSLSITQDLDGDDGSDPTPMPLEFETRTRESA
VTRNNNSVSSREHRSPTSTGRGRRNADSPVPEPATSLRRATAGPWS
PSSYLTIDLTDEDHTGPTSSSTYARLADSOOGMESPEAOAGSDSESO
DKKEFIGDLVNGKIKGFWPAMVSWKATSRKQMSGMRVOWFGDKSEVSAK
LVALGIFSOHMLATFNKLVSRKAMHALEKARVAPKTPFSSGDSLEQKLPME
MAHGKRPTEIGLAPNNTOPVNRKSVRRAGSRKLESKRKYNNKRRRTADSTSY
CPAPKRLKTCNNNGKRDGDDOSREOMASDVANNKSLSDCSLSCGRRNVSHLP
EGGLQTCRDLLELYMDDGYOYCTVCCGRELDCSTSCRCFEVCELEVLV
GTGTAAEKLQEPSCYMLCPORCHGVLRKRMNVRLOAFSTGLEVAPKLYA
IPARRRPTRVSLDFGIATGYLVKELGKIVKRVASVECESIAVTVHEGNITV
VNDNRITKKNIEWGPDLVIGSPNDLSNVNPARKGLYEGSTDFEFYHLNYS
RPEGDDREFPMFENVAMVGDGRDISRFELKLVNIDAIKVSANARAYEWGNP
GNRPVYASKNDKLELDCLERYNRIAKLVQITTKKSNISIGKNNLFPVNMKGD
VLWCTELERIFGFVHYTDVSNMGRAROKRLGRSWSVEVIRHLEAPLKDFACE"

gene
CDS

BASE COUNT 1041 a 1083 c 1096 g 925 t
ORIGIN

Query Match 43.2%; Score 1812.2; DB 9; Length 4145;
Best Local Similarity 76.2%; Pred. No. 0;
Matches 2532; Conservative 0; Mismatches 653; Indels 138; Gaps 19;
OY 251 GCCCTCTTTCAGAACATGAAGAGAGACACACATCTGAATGAGAGAGGTGCC 310
Db 97 GCCCTCGCGGAGAAACATGAAGAGAGACACACATCTGAATGAGAGAGGTGCC 156
OY 311 AGCGGGTATGAGAGTGCATTTATGTTAGGAACTTCACTGACAGCTCTGACAGC 370
Db 157 GCGGGAGAGAGAGTGCATTTATGTTAGGAACTTCACTGACAGCTCTGACAGC 212
OY 371 AAGATGCTCCCTACCGCCAGTCTTGAGAGCAATCTGCACAGCCAGTCTGACAGCA 430
Db 213 -----CTGCCCCCAATCTGAGAGGCTA-----TCCGACCCCG 246
OY 431 GAGACAGAGCGCCAGAGTCAAGCTCCGCGCTGCTTAAGAGAGGTCTGACAGCTTCTG 490
Db 247 GAGATCAGAGCGCCAGAGTCAAGCTCCGCGCTGCTTAAGAGAGGTCTGACAGCTTCTG 306
OY 491 AATTACACGACAGACATGACAGAGATGAGACAGAGATGATGATGATGATGATGAT 550
Db 307 AGCTACACACAGAGACTTGCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 354
OY 551 GGCTCTGATATTC---TAATGCCAAGCTCACCGCTGAGACCAAGACACAGAGAGAGAG 607
Db 355 GGCTCTGACACCCAGTCAATGATGACCAAGAGCTTCTGCG-----GGAAACAGAGAGCTTCTG 405
OY 608 TCTGAAAGCGCGGCTGTCGACACCGACATAGCAATGAGAGAGCTTCTGAGAGAGCA 667
Db 406 TCGAAGAGCGCGGCTGTCGACACCGACATAGCAATGAGAGAGCTTCTGAGAGAGCA 465
OY 668 AGAGCTCTCCCAAGATACACCGAGTCCGACAGGCGCCGACATCTGACAGAGAGTCT 727
Db 466 AGAGCTCTCCCAAGATACACCGAGTCCGACAGGCGCCGACATCTGACAGAGAGTCT 525
OY 728 GTGAGATTTCCGCTGACAGAGTCTCGAGAGCTCGAGATGCTTCTGACAGAGAGCA 787
Db 526 GTGAGATTTCCGCTGACAGAGTCTCGAGAGCTCGAGATGCTTCTGACAGAGAGCA 585
OY 788 TGGTCATCCCTGCGAGGCTGACCTTCATGAGAAAGT----- 825
Db 586 TGGCTCTCCCTGCGAGGCTGACCTTCATGAGAAAGT----- 825
OY 826 -----GACACTAAGAGCGTCACTGACCTTCACTGACCTTCACTGACCTTCACTG 880
Db 646 CATGGAGCGCCCGAG 705
OY 881 GAGGTATGATGATGACACACAGTGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 940
Db 706 GAGGTATGATGATGACACACAGTGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 765
OY 941 CAGCATGATTAAGAGTGTGAATAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1000
Db 766 CAGCATGATTAAGAGTGTGAATAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 825
OY 1001 TGGTGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1060
Db 826 TGGTGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 885
OY 1061 ATGCGCTGGGTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1120
Db 886 ATGCGCTGGGTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 945
OY 1121 GTGGCTGCTGGGTCTTACACACAGTGTGAATGATGATGATGATGATGATGATGATG 1180
Db 946 GTGGCTGCTGGGTCTTACACACAGTGTGAATGATGATGATGATGATGATGATGATG 1005
OY 1181 TATGAGAGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1240
Db 1006 TATGAGAGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1065

QY	1241	TCACGAGCTCTCGAGAGTCTACTGTGAGGACCAGCTGAAGCCATGCTGGAGTGGGCCAC	1300
Db	1066	CCGACGAGCCCTGGAGACTCATTTGGAGAGCACGCTGAAGCCATGTTGGTGGGCCAC	1125
QY	1301	GGTGGCTTCAAGACCTTACTGGGATGGAGGGCCCTCAACCCCAACAGAACCAACGTGGTT	1360
Db	1126	GGGGGCTTCANAGCCCACTGGGATTCGAGGGCCTCAACCCCAACACAGCAACCACTGGTT	1185
QY	1361	AATAAGTCGAAGGTGCGCTCGTTTCAGACAGTGAAGAACTTAGAACCCAGAGACGGAGAAC	1420
Db	1186	AATAAGTCGAAGGTGCGCTCGTGCAGAGGAGTAGAATAATTAAGATCAAGAAATACGAGAC	1245
QY	1421	AAAAGTCGAAGACGCACACACCACTGACTCTGCTCTCTG---AGTCCCCCAACCAAG	1477
Db	1246	AAGACTCGAAGACGCACAGCTGACAGCTCACCCACTCTGACTCTCTCCCCGACCAAG	1305
QY	1478	CGCCTCAAGCAAAATAGCTAT---GGGGGGAAGACCGAGGGGGAGTAGAGACAGCGA	1534
Db	1306	CGCCTCAAGCAAAATAGCTATACACAGCGCAAAACCGAGGGGATGAAGATCAAGCGA	1365
QY	1535	GAACGAGATGGCTTTGAAGTACACACACAAACAAAGGGCAATCTGGAAGAAGCGCTGTTTGC	1594
Db	1366	GAACAATATGGCTTAGATGTTGGCCACACAGACAGACGCTCGAAGATGGCTGTTTGTCT	1425
QY	1595	TGTGAAAAAGAAACCTGTGTCCYCCACCCCTCTTTGAAGGTGGGCTCTGTACAGT	1654
Db	1426	TGTGGCAGAAAAACCCGCTGTCTTCACCCCTCTCTTTGAGGGGGGCTCTGTACAGCA	1485
QY	1655	TGCCGGATTCGCTTCCCTAGAGCTCTTTACATGATATATAGAGACGGCTATCACTCTAC	1714
Db	1486	TGCCGGATTCGCTTCCCTAGAGCTCTTTTACATGATATATGACATGCTATCACTCTTAC	1545
QY	1715	TGCACCTGTGCTGTAGAGGGCCGTGAACCTGCTGTGTGACACTATACACAAAGCTGTGCAGA	1774
Db	1546	TGCACCTGTGCTGTAGAGGGCCGTGAACCTGCTGTGTGACACTATACACAAAGCTGTGCAG	1605
QY	1775	TGCTTGTGTGAGAGTCTGTGAGAGTGTGTGGGCGCAGCAGACGCTGAGAGTCCAG	1834
Db	1606	TGTTTCTGTGTGAGAGTGTGTGAGAGTGTGTGGGCGCAGCAGCAGCGCCGAGCAAG	1685
QY	1835	CTGACGAGAACCTTGAGACTGCTATATATGTCCTCCCTCAGCCCTGCCATGGGGTCTCTCGA	1894
Db	1666	CTTGAGAGGACCTTGAGAGTGTCTATGTCCTCCCGAGCGCTGTCAATGGGTCCTCGCG	1725
QY	1895	CGCAGAGAAATTTGGAACATCGCGCTCGCAACACTCTTCACTACTGATCCGTGACCTGGAA	1954
Db	1726	CGCGGAGAGACTGGAACGTGCGCTTGCAGGCCCTTCTTCACCACTGACACGGGCTTGA	1785
QY	1955	GAATTTGAGCCACCCAAAGTTGTACCCAGCAATTCCTGACAGCCAAAGAGAGGCCATTAGA	2014
Db	1786	TAC---GAGGCCCCCAAGCTGTACCTGTGCCATTCGCCGAGGCCGAGGGGCCATTGCA	1842
QY	2015	GTCCTGTCTCTGTTTATGATGAAATTGCAACGGGGTACTTGGTCTCAAGAGAGTTGGTATT	2074
Db	1843	GTCCTGTCTCTGTTTATGATGAAATTGCAACGGGGTACTTGGTCTCAAGAGAGTTGGGCTA	1902
QY	2075	AAAGGGAAGAAAGTACATTTGACCTCGAAGTGTGTGACAGGCTCATGCGGTGGGAAGCTT	2134
Db	1903	AAGGTAGAAAGTACATTTGACCTCGAAGTGTGTGACAGGCTCATGCTGTGTGGGAAGCTG	1962
QY	2135	AAGCATGAAGCCAGATCAATAATGTCAATGAGCTCCGGAATAATCACAAAGAAATAATT	2194
Db	1963	AAGCATGAAGCCAGATCAATAATGTCAATGAGCTCCGGAATAATCACAAAGAAATAATT	2022
QY	2195	GAAGAGTGGGGCCCTTGACCTTGGTATTTGGTGAAGCCCATGCAATGATCTCTTAC	2254
Db	2023	GAAGAGTGGGGCCCTTATTTGATTTGGCGGAAGCCCATGCAAGCAATCTCTCAAAAT	2082
QY	2255	GTCAATCTCGCCCAAGGTTTATATAGAGGGCAAGGAAGGCTTCTTGATGATTTTAC	2314
Db	2083	GTGAATCTCGCCCAAGGTTTATATAGAGGGTACAGGCGGCTTCTTGATGATTTTAC	2142
QY	2315	CACTTGTGAATTATCCCGCCCAAGAGGGGCACAAACGCTTCATTTCTTGATGTTC	2374

[illegible]

Db	3204	CAGAAAG-----AGAAAAATGTTGTAATATGTC-----TTTACCGCGGACATTCGCCCTTGCTCA	3255
OY	3441	GGTCAACAGGCGTGAAGTCTGCTGTCG-----TCTTGTAGACATTTCCAGCATGATGATG	3493
Db	3256	AATCAACAGGCGTGTGATGTGCACGGACGACCATTAAGATGATTTTCCAAATGATGATGAT	3315
OY	3494	TCACGACGAGTCATCATCCACAC	3516
Db	3316	TCACGACGAGTCATCCATCATC	3338
RESULT 15			
LOCUS	AF176228	4267 bp	mRNA linear PRI 26-OCT-1999
DEFINITION	Human sapiens DNA cytosine-5' methyltransferase 3B (DNMT3B) mRNA,		
ACCESSION	AF176228		
VERSION	AF176228		
KEYWORDS	complete cds.		
ORGANISM	AF176228.1 GI:6118091		
SOURCE	Human sapiens.		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1. (bases 1 to 4267)		
TITLE	Xu, G.-L., Bestor, T. H., Bourc'h, D., Hsieh, C.-L., Tommerup, N.,		
UTHERS	Bugge, M., Hulten, M., Ou, X., Russo, J. J. and Viegas-Pequignot, E.		
FEATURES	Chromosome instability and immunodeficiency syndrome caused by		
FEATURES	mutations in a DNA methyltransferase gene		
SOURCE	Nature (1999) in press		
JOURNAL	2 (bases 1 to 4267)		
REFERENCE	Xu, G.-L. and Bestor, T. H.		
AUTHORS	Direct Submission		
TITLE	Submitted (06-AUG-1999) Genetics and Development, Columbia		
JOURNAL	University, 701 West 168 St., New York, NY 10032, USA		
LOCUS	Location/Qualifiers		
FEATURES	1..4267		
SOURCE	/organism="Homo sapiens"		
FEATURES	/db_xref="taxon:9606"		
SOURCE	/chromosome="20"		
FEATURES	/map="20q11-q13"		
SOURCE	1..4267		
FEATURES	/gene="DNMT3B"		
SOURCE	265..2802		
FEATURES	/note="mutated in human chromosome instability and		
SOURCE	immunodeficiency disease known as ICF syndrome; contains		
FEATURES	alternative 5' exon 1B"		
SOURCE	/codon_start=1		
FEATURES	/product="DNA cytosine-5' methyltransferase 3B"		
SOURCE	/protein_id="AF04015.1"		
FEATURES	/translation="MESPPEPSSLESKSKDTRHLNGEEDDGGEDSLIYNACSSOSS		
SOURCE	DSPPILAEATPEIRNGKSSSKLRKSRVSLSLTQSDGDDGDDGDDSDPPVKKPDK		
FEATURES	REFRTRBESPAKRRNNNSVSKSRKRRVSSSTQSDGDDGDDGDDSDPPVKKPDK		
SOURCE	KATASAGTSPFSPSSITLIDLDIEDTHGTFSSSPKATKATKATKATKATKATKATKATKAT		
FEATURES	ADSDGDSSTEDGKEFGLIDLVNKKIKKESSTKATKATKATKATKATKATKATKATKATKAT		
SOURCE	GDGKFSVSKEDKLVYKGLFSDFHMLKATKATKATKATKATKATKATKATKATKATKAT		
FEATURES	DSLDSDQKPKLEKLVNKKIKKESSTKATKATKATKATKATKATKATKATKATKATKAT		
SOURCE	INCLNNKKGDDSD		
FEATURES	KDRLEFLTHDDDDGGLSCGCGSSKATKATKATKATKATKATKATKATKATKATKAT		
SOURCE	KLOPEWCHICLQKCHQHLKRNKNNVYELDPPSDNCELYEPKLYATATTAARRRR		
FEATURES	IRVSLSDGATGILGLLELRLKRNKNNVYELDPPSDNCELYEPKLYATATTAARRRR		
SOURCE	KKNLEFGSPDYLIGSSKNDLVNKKIKKESSTKATKATKATKATKATKATKATKATKAT		
FEATURES	PFKMFENYVAKVGGKDKSRITDELKATKATKATKATKATKATKATKATKATKATKAT		
SOURCE	SKNKKLQDQCSKNNKLNKLNKLVDTLTSNSIKGKNQDLPPVYMGKEDVIMCTELE		
FEATURES	SKNKKLQDQCSKNNKLNKLNKLVDTLTSNSIKGKNQDLPPVYMGKEDVIMCTELE		
SOURCE	BASE COUNT 1069 a 1098 c 1147 g 953 t		
ORIGIN			
Query Match	40.5%	Score 1697.2	DB 9: Length 4267;
Best Local Similarity	74.7%	Fed No. 0:	
Matches 2495; Conservative	0;	Mismatches 648;	Indels 199; Gaps 21

OY	233	TTTGGGGACCAATTCGAGGGGCTTCCTTTTC-AGGAAACAATSAAGGGAGACAGACATCT	291
Db	264	CATGGAACCAACATCTCTGAGCCTCCACAGCTTGGAAAGCATGAAGGGAGACACAGGACATCT	323
OY	292	GAATGAAGAAGAGGGGTGGCCAGCGGGATATGAGAGTGCAATTAATCTTAAATGGACATTCAG	351
Db	324	CAATGCGAGAGGGAGAGCGCGCGGGAGGGAAAGCTACATCTCTGTACACGGGGGCTGTGAG	383
OY	352	TGAGCATCTCTCAGACAGGAAGATGCTCTCTACCCCTCCAGCTCTGGAGGCAATCTGCAC	411
Db	384	CGACCAATCTCTCGA-----CTGGCGCCCAATCTCTGAGAGCTA-----	421
OY	412	AGAGCCAGTGTGCACACCGAGACAGAGCGCCGACAGCTCAACCTCCCGGCTCTAAAG	471
Db	422	-----TCCGACCCCGGGAGATCACAGGACCGGAATCAGAGCTCCGCACTCTCCAAAG	473
OY	472	GGAAGTCTCAGCCTTCCTTGAAATATACAGCAAGGATGAGAGATGAGAGACAGATGA	531
Db	474	GGAAGTCTCAGCTCTGCTTACCTTACACAGAGATCTGACAGGGGATATGGGCA-----	525
OY	523	TGAAGTGAATGGAGGAATGGCTCTGTGAATATC-----TATGGCAAGCTGACCCGCTGAGAC	588
Db	526	-----GGGGAAGATGGGATGGCTCTGACACCCCATGATGCGCAAAAGCTCTCG-----	575
OY	588	CAAGAGACAGGAGAGAGGAGCTGTGAAGAACCGGGCTGCGGAACCCGACATATGCAATGGGAC	648
Db	576	---GGAAACCAAGGACTGTCTCAAGAAACCCAGCTGTCCGAACCTGAAATATACAAAGTGT	632
OY	649	CTCCAGCTTGGAGAGGCAAGAGCCCTCCCGCAAGAAATCAACCCGAGTGTGGCAGGGCCGCA	708
Db	633	CTCCAGCCGGGAAGGACAGAGCCCTTCCCGACGTTCCACCCGAGGCGCGAGGGCCGCA	692
OY	709	CCATGTCTCAGAGTACCTCGTGTGAGATTCCGGCTACACAGGTCTGTGAGAGCTGTGAGATC	768
Db	693	CCATGTGAGAGATCTCCCGGTGTGAGATTCCGGCTACACAGGTCTGTGAGAGGCGGGCAAC	752
OY	769	GCTCTCAGCAAGACGCCATGGTATCCTCCCTGTGCAGAGCTCACTCA-----	815
Db	753	AGCATCGGCGAGAAAGCCATGGCCGTCCCTCCAGCTCTTAACCTTACATGACCTCAC	812
OY	816	-----TGGAAGAGTGAACCTTAAGAGAGCTGAGTACCCCATCAATGACTT	861
Db	813	AGACGACACAGAGAGACACATAGGAGAGCCCCAGACACAGAGTACCCCTACGCCGCT	872
OY	862	GAGCCAGGATCGAGATCAGAGAGGATATGATTAACACACAGGTGATGCAGAGACAGAGA	921
Db	873	AGCCCCAGGACAGCCACAGAGGGGCGCATGGAGTCCCGCAGGTGAGAGCACAGCTGGAGGA	932
OY	922	TGAGACAGCAGCAGATATCAGAGATGATTAAGAGATTGGATAGTGAAGTCTGTGGGG	981
Db	933	TGAGACAGCTTCAGAGATATCAGAGATGGGAAGAGATTGGAATAGGGAGCTCTGTGGGG	992
OY	982	AAAGATCAAGGGCTTCTCTCTGGTGGCTCCCATGTGTGTCTGGAAAGCCACTCCAA	1041
Db	993	AAAGATCAAGGGCTTCTCTCTGGTGGCCCGCATGTGTGTCTGGAAAGCCACTCCAA	1052
OY	1042	GCGACAGGCCATGCCCGGAATGCGCTGGGTGACAGTGCTTGGTGATGGCAATTTTCGA	1101
Db	1053	GCGACAGGCTATGTGTGAGATGCGGTGGGTGCAGTGCTTGGCAATGGCAATTTTCGGA	1112
OY	1102	GATCTCTGTGACAAACTGTGTGGCTGTGGGGCTGTTCAGCCAGCACTTAAATGTGGCTAC	1161
Db	1113	GGTCTCTCAGACAAACTGTGTGGCACTGTGGGGCTGTTCAGCCAGCACTTAAATTTGGCCAC	1172
OY	1162	CTTTCATAAGCTGTGTTCTTATAGGAAGCATGTACACACTCTGGAAGAAACCGAGGT	1221
Db	1173	CTTTCATAAGCTGTCTCTCTATATCGAAAGCATGTATACATGCTCTGGAAGAAACCTAGGT	1232
OY	1222	TGAGCTGTGCAAGACTTCTTCAGCATCTCTGAGAGTGTACATGAGAGACAGCATGAAAGC	1281
Db	1233	GCGAGCTGTGCAAGACTCTCTCCAGAGCTCTGGAGACTCATTTGAGAGACACAGCTGAAGC	1292
OY	1282	CATGCTGAGATGGGCCACAGGTGTCTCAAGCTACTGTGGATCGAGAGGCTCAAAACCA	1341

[illegible]

Dp	2310	SCCGTCTTTCGATGTTTTGAGAAATGTTGTAGCCATGAAAGTTGGCCACAAGAGGACAT	2309
Qy	2416	CTCAAGATTCCTGGCAGGTAAACCCAGTAGATGATGCATCCATCAAGAGTGTCTGCTCTCA	2475
Dp	2370	CTCAAGGTTCTCTGGAGTGTAAATCCAGTGATGATTTGATGCCATCAAGTTTCTGCTGTCA	2429
Qy	2476	CAGGGCCCGGACTCTTCTGGGGTAACTTACCCTGGAATGACAGGCGCCGTGATGGCTTAAA	2535
Dp	2430	CAGGGCCCGAATCTTCTGGGGCAACCTTACCCTGGGATGACAGGCGCCGTGATGATCAAAA	2489
Qy	2536	GAATGATTAAGCTCGAGTGCAGGAGCTGCTGGAGTCAATGAGGACAGCAAAAGTTAAAGAA	2595
Dp	2490	GAATGATTAAGCTCGAGTGCAGGAGCTGTTGGATTACAAATGATATGCAAGTTAAAGAA	2549
Qy	2596	AGTGCAACAATTAACACCAAGTAGAATCCATCATGACAGAGGCCAAAAACAGCTTTTCCC	2655
Dp	2550	AGTACAGACAAATTAACCAACCAAGTAGAATCCATCATGACAGAGGCCAAAAACAGCTTTCCC	2609
Qy	2656	TGTACTCATTAATGGCAAGACAGACTTTTGTGTGCACTGAGCTCGAAAGATCTTGGG	2715
Dp	2610	TGTTGTCTGTAATGGCAAGAAAGATGTTTGTGTGCACTGAGCTCGAAAGATCTTGGG	2669
Qy	2716	CTTCCGCTCACTAACACAGGAGCTGTCCAAATGAGGCGGGGCGCCGTCAGAGAGCTCT	2775
Dp	2670	CTTTCCTGTGCACTAACAGAGCTGTCCAAATGAGGCGGGGCGCCGTCAGAGAGCTCT	2729
Qy	2776	GGGCAAGTCTCTGGAGGTGTACCGGTTCATTCAGACACTGTGTTGGCCCTTGAAGAGCTACTT	2835
Dp	2730	GGAAAGGTCTTGGAGCGTGCCTGTTCATTCGACACTGTGTGCGCCCTGTGAAGAGCTACTT	2789
Qy	2836	TGCTGTGTAATAGTTCTTACCCAGAGACTGGAGACTCTCG-----GTCAAGCCAGTGC	2888
Dp	2790	TGCATGTGAATAGTTCTTACCCAGAGCGCCCAACCCCACTGGGGTGTGTGGCAGAGCAGAGAC	2849
Qy	2889	CCAGAGTCAACCCCTCCCTCCGAAGGAGACCTCACCTGCCCCCTTTTAAAGTCACCTGTGGGG	2948
Dp	2850	CCAGAGTGTGATTTCTCTGAAGGGA--TCCCAAGGCCCTGCTCTTCTCAGACTGTGGG	2907
Qy	2949	GCCTCACATCACTGTACCTCAGCTTCTTCTGCTCAGTGGGAGGAGACCTCCCTGGGCGCT	3008
Dp	2908	TCAT-----ACCGTGAACCTCAGTTTCCCTTGTGTGATGGGGGAGACCACTGACCTCT	2963
Qy	3009	TGCAGGGGAGCCC--CGGTGCTCCCTCCGTGTGCACAGCTTAGACCTGCTTAGAGT	3066
Dp	2964	TGCAGGGGTACCTGAGGGTGCCTGCTCTTGTGCACAAATCAGACCTGGCTGTGAGAGC	3023
Qy	3067	AGCCCGGAGGTGCTCATGTTTCTTAACTCCCTGAAACTTTAAACTTGAATAGTATAGTA	3126
Dp	3024	AGCCTAACAGGTGCTCATTTTCTTCTTAACTTAACTTAACTTGAATAGTATAGTATAGTA	3083
Qy	3127	AGATGGC---TTCCTTTTACCCTCCCTGAGTTTATCACTC--AGAGTGAATGGCTAAGATA	3181
Dp	3084	ACGTGGCTTTTTTTTTTCCCTCTCTGGGTGTACCACTCAGAGAAACAGTGGCTAAGATA	3143
Qy	3182	CCAAAAAACAACAAAAACAGAAACAAAAAACAATAAACTCAACAGCTCTCTTAA	3241
Dp	3144	CCAAACACACAGTGCAGACCTCTCCAACTCAGGTTAATGCTGAAAAATCAATCAACAG	3203
Qy	3242	TACTCAGGTATGCTGCAAAATCACTGAGATTTTGTTTTAAAGTAAACCGTGTCTCAC	3301
Dp	3204	ACAGTTATTCAGAGTTTA-----TTTGTAAACTGGCTATCTCTCTGCT	3250
Qy	3302	ATTGTCGTGGAGATGCTATTTGATGTTGGGCTCAGATGAGACAGGTCAGAGGGCCAAAA	3361
Dp	3251	GTTTACACAGCTGTGCAAGTTGTAGGACGTAGCTACAGGACATTTTAAAGGCCCAAGAT	3310
Qy	3362	AAAATTCCTCCTCCGCCAGGAGTATTTGAAGTGTATTGTTATGCTTAACTCTTCT	3421
Dp	3311	CGTTTTCCTCCAGGGCAAGAGAA--AGAAAATGTTGATATATGCTCT---TTTACCC	3362
Qy	3422	GGCACTTCCCTTGGCTTGGTACAAAGGCGTGAAGTCTCTGTG-----TCTTGAAG	3474
Dp	3363	GGCACTTCCCTTGGCTTAAATTAACAAAGGCGTGGAGTGTGCACGGGACCTATTAGAGTAT	3422

Tue Jul 22 16:21:53 2003

us-09-720-086-2.rge

Page 36

Oy 3475 ATTTCCACGATGATGTCTACAGGGATGCATCACCACC 3516
 | | | | | | | | | |
Db 3423 TTCACAATGATGATTTCACGAGGGATGACGTCATCATC 3464

Search completed: July 17, 2003, 07:37:39
Job time : 7065.06 secs

RESULT 1
 AF068626
 LOCUS
 DEFINITION Mus musculus DNA cytosine-5 methyltransferase 3b1 (Dnmt3b) mRNA, 4195 bp
 ACCESSION AF068626
 VERSION AF068626
 KEYWORDS alternatively spliced, complete cds.
 SOURCE AF068626.2 GI:6449469
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Cloning and characterization of a family of novel mammalian DNA
 Okano, M., Xie, S. and Li, E.
 1 (bases 1 to 4195)
 REFERENCE
 AUTHORS
 TITLE

THIS PAGE BLANK (USPTO)

(cytosine-5) methyltransferases
JOURNAL Nat. Genet. 19 (3), 219-220 (1998)
MEDLINE 98324766
PUBMED 9662389
REFERENCE
2 (bases 1 to 4195)
Xie, S., Okano, M. and Li, E.
Direct Submission
Submitted (28-MAY-1998) CVRC, Mass. Gen. Hospital, 149 13th Street,
Charlestown, MA 02129, USA
3 (bases 1 to 4195)
Okano, M., Chijiwa, T., Sasaki, H. and Li, E.
Direct Submission
Submitted (04-NOV-1999) CVRC, Mass. Gen. Hospital, 149 13th Street,
Charlestown, MA 02129, USA
REMARK
Sequence update by submitter
On Nov 18, 1999 this sequence version replaced g1:3327979.
COMMENT
FEATURES
Source
Location/Qualifiers
1..4195
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="2"
/map="84.0 cM"
1..4195
/gene="Dnmt3b"
1..4195
/gene="Dnmt3b"
/note="similar to EST sequences deposited in Genbank
Accession Numbers A116694, A119979, A117277, A1210568,
AA407106, and AA575617"
269..2848
/gene="Dnmt3b"
/function="de novo DNA methylation"
/note="alternatively spliced product; contains Cys-rich
region; C-terminal region is similar to corresponding
region of DNA cytosine-5 methyltransferase 3A"
/codon_start=1
/product="DNA cytosine-5 methyltransferase 3b1"
/protein_id="AAC40178.2"
/db_xref="GI:6449470"
/translation="MKGDSRLNEEGASGEYECITVNGNFSDSSDTRKDA
VICTPCTPETERGRSSRLSKREVSLNVTOMTDGDRDDVDNDGSDIMPK
LRTREKDTRESSEPAVTRSHNGTSLEORASPRITRGGRHNDVEVEPATR
SRRRASASASPSSPASVDPMEEVSTKSTSPVDSIQGDGDDMTTVDSESD
GDSREYDDKKEFGIDLWCKIKGFSMPKAVVSKATSKQAMPGMVQWFGDKF
SEISDKRLVALGLSOFHNLATFNKLVSRYKAMYHLEKARVAGKTFSSSGFLED
OLKPLERAHGGEKPTGIEGLPKPKKQPVVSKVRSRSDSNLEPRRENKSRRTTN
DASASESPPRKRLKNTSYGKRDGEDSESRMASSEVTNNKNTLEDCLSGKKNPS
FHPLEGGICSDRLELFYNYDEGQYSCVCCBGRLLCSNTECCRCVCEC
LEVAVGATADADLOEPMSCVCLPQCHVLRKRDMMNRLLDFTTDDLEEFER
PKLYPAIPAKRRIRIVLSLFDGIATGILVLEKLGITVEYIASEVCAESIAGTVH
EGOIKYNDVYKTRKKNIBEMGPDVIGSPNDLSNVNPARGLGEGRLFEFY
HLNLTYPKEDNRFTEFENNVAMVNDKDISRLACNPVIMDAIKYSAHARY
FKGNLPGMNRVASKNDKLELDQCLEFSESTAKLKVTQITTSNLSROCKNOLFVY
MNGKDLVWCTELERIPFPAHYTDVSNMGRGAROKLGLSMSPVIRHLPAPLKDIF
ACE"

gene
misc-feature
CDS

BASE COUNT 1061 a 1059 c 1125 g 950 t
ORIGIN

Query Match 100.0%; Score 4195; DB 10; Length 4195;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GAATTCGGGGCGGGGCTTAACGGCCCAAGTAAAGTAGCCAGCAGATCGCGCCGGA 60
1 GAATTCGGGGCGGGGCTTAACGGCCCAAGTAAAGTAGCCAGCAGATCGCGCCGGA 60
61 GATTTCGGAACCGAAGTCCGGCGCCCGCGCGGAGAGACCGCGCGCGATCGCGGC 120
61 GATTTCGGAACCGAAGTCCGGCGCCCGCGCGGAGAGACCGCGCGCGATCGCGGC 120
121 GCCGCGCTACAGCCAGCCTACAGACAGCGCCCGCTGAGCGCTTGCCAGACCTTGAAGAC 180
121 GCCGCGCTACAGCCAGCCTACAGACAGCGCCCGCTGAGCGCTTGCCAGACCTTGAAGAC 180

181 TCAGTATATACCTTTCAGACAGCGGGATCTCCCTCCCATCATAGTGCCTGGGAC 240
181 TCAGTATATACCTTTCAGACAGCGGGATCTCCCTCCCATCATAGTGCCTGGGAC 240
241 CAAATCCAGGCGCTTCTTTCAGGAACAATGAAGAGACAGCAGATCGATGATGAAGA 300
241 CAAATCCAGGCGCTTCTTTCAGGAACAATGAAGAGACAGCAGATCGATGATGAAGA 300
301 AGAGGTCGACAGGCGGATAGAGAGTCAATTCCTTAATGGAATTCAGTACAGCAGTC 360
301 AGAGGTCGACAGGCGGATAGAGAGTCAATTCCTTAATGGAATTCAGTACAGCAGTC 360
361 CTCAGACAGCAAGAGTCTCCCTACACCCCAAGCTTGGAGGCAATTCGACAGACGCA 420
361 CTCAGACAGCAAGAGTCTCCCTACACCCCAAGCTTGGAGGCAATTCGACAGACGCA 420
421 CTCAGACAGCAAGAGTCTCCCTACACCCCAAGCTTGGAGGCAATTCGACAGACGCA 480
421 CTCAGACAGCAAGAGTCTCCCTACACCCCAAGCTTGGAGGCAATTCGACAGACGCA 480
481 CAGCCTTCTGAATTAACGACAGACATGACAGAGATGAGAGACAGATGATGAAGTAA 540
481 CAGCCTTCTGAATTAACGACAGACATGACAGAGATGAGAGACAGATGATGAAGTAA 540
541 TGATGGGAATGCTCTGATATCTTAATGCCAAGCTCCCGCTGATGAGACAGACAGCAG 600
541 TGATGGGAATGCTCTGATATCTTAATGCCAAGCTCCCGCTGATGAGACAGACAGCAG 600
601 GACGCGCTTGAAGACCGCGCTGTCGCAACCCGACATAGATGAGGACCTCCAGCTTGA 660
601 GACGCGCTTGAAGACCGCGCTGTCGCAACCCGACATAGATGAGGACCTCCAGCTTGA 660
661 GAGGCAAGAGAGCTCCCGCAATACCCGAGAGTGGCAGGCGCCACCATGTGCAGGA 720
661 GAGGCAAGAGAGCTCCCGCAATACCCGAGAGTGGCAGGCGCCACCATGTGCAGGA 720
721 GTACCTGTGAGATTCCGGCTACAGAGTCTCGAGAGTGGAGATGATGATGATGATGAT 780
721 GTACCTGTGAGATTCCGGCTACAGAGTCTCGAGAGTGGAGATGATGATGATGATGAT 780
781 CAGGCAAGAGAGCTCCCGCAATACCCGAGAGTGGCAGGCGCCACCATGTGCAGGA 840
781 CAGGCAAGAGAGCTCCCGCAATACCCGAGAGTGGCAGGCGCCACCATGTGCAGGA 840
841 CAGTACCCCATCACTTGAATGACCTGACCCAGAGATGAGAGTGGATGATGATGATGAT 900
841 CAGTACCCCATCACTTGAATGACCTGACCCAGAGATGAGAGTGGATGATGATGATGAT 900
901 GGTGATGACAGAGAGAGATGAGAGACAGACAGATGATGATGATGATGATGATGAT 960
901 GGTGATGACAGAGAGAGATGAGAGACAGACAGATGATGATGATGATGATGATGAT 960
961 AATAGTGAGCTGCTGCTGGGGGAAAGATCAAGGCTCTCTGCTGCTGCTGCTGCTGCT 1020
961 AATAGTGAGCTGCTGCTGGGGGAAAGATCAAGGCTCTCTGCTGCTGCTGCTGCTGCT 1020
1021 GTCTGGAAGAGCAGCTCCAGAGGAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
1021 GTCTGGAAGAGCAGCTCCAGAGGAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
1081 TGGTATGCAAGATTCTGAGATCTCTGCAACAATGCTGCTGCTGCTGCTGCTGCTGCT 1140
1081 TGGTATGCAAGATTCTGAGATCTCTGCAACAATGCTGCTGCTGCTGCTGCTGCTGCT 1140
1141 CCAGCACTTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
1141 CCAGCACTTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
1201 CACTGTGGAAGAGCAGAGGCTGAGAGTGGCAAGACCTTCTCAGACAGTCTGAGAGATC 1260
1201 CACTGTGGAAGAGCAGAGGCTGAGAGTGGCAAGACCTTCTCAGACAGTCTGAGAGATC 1260

THIS PAGE BLANK (USPTO)

OY	1261	ACTGAGACAGCTGAAGCCCATGCTGAGTGGGCCACGCTGCTCAAGCCTACTGG	1320
DB	1261	ACTGAGACAGCTGAAGCCCATGCTGAGTGGGCCACGCTGCTCAAGCCTACTGG	1320
OY	1321	GATGAGGAGGCTCAACCCCAACAGAACACAGTGGTTAATAGTGAAGTGGCTCG	1380
DB	1321	GATGAGGAGGCTCAACCCCAACAGAACACAGTGGTTAATAGTGAAGTGGCTCG	1380
OY	1381	TTGACAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	1440
DB	1381	TTGACAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	1440
OY	1441	CAATGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1500
DB	1441	CAATGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1500
OY	1501	CGGGAAGAGCCGAGGAGGAGTGAAGAGCCGAGAACGATGGCTTCTGAAGTCAACA	1560
DB	1501	CGGGAAGAGCCGAGGAGGAGTGAAGAGCCGAGAACGATGGCTTCTGAAGTCAACA	1560
OY	1561	CAACAAGGCAATCTGAAGACCGCTGTTTCTCTGTGGAAGAAAGAAACCTGTCTT	1620
DB	1561	CAACAAGGCAATCTGAAGACCGCTGTTTCTCTGTGGAAGAAAGAAACCTGTCTT	1620
OY	1621	CCACCCCTCTTTGAGGGGTGCTGTCAAGTGGCGGGATGCTGCTGAGAGTCTCT	1680
DB	1621	CCACCCCTCTTTGAGGGGTGCTGTCAAGTGGCGGGATGCTGCTGAGAGTCTCT	1680
OY	1681	CTACATGTATGATGAGAGCGCTATCAGTCTCTACTGACCGCTGCTGAGAGCCGTA	1740
DB	1681	CTACATGTATGATGAGAGCGCTATCAGTCTCTACTGACCGCTGCTGAGAGCCGTA	1740
OY	1741	ACTGCTGTGTGCAATACACAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1800
DB	1741	ACTGCTGTGTGCAATACACAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1800
OY	1801	GCTGTGGGGCGGAGCAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	1860
DB	1801	GCTGTGGGGCGGAGCAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	1860
OY	1861	GTCCTCTCCCTCAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1920
DB	1861	GTCCTCTCCCTCAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1920
OY	1921	GCAAGACTTCTTCACTACTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG	1980
DB	1921	GCAAGACTTCTTCACTACTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG	1980
OY	1981	AGCAATTCCTGAGCCAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	2040
DB	1981	AGCAATTCCTGAGCCAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	2040
OY	2041	AACGGGCTACTTGTGCTCAGAGAGTGGGTATTAAGTGAAGTGAAGTGAAGTGAAG	2100
DB	2041	AACGGGCTACTTGTGCTCAGAGAGTGGGTATTAAGTGAAGTGAAGTGAAGTGAAG	2100
OY	2101	AGTCTGTGAGAGTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2160
DB	2101	AGTCTGTGAGAGTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2160
OY	2161	CATGAGCTGCGGAAATATCAGCAAGAAATATTAAGTGAAGTGAAGTGAAGTGAAG	2220
DB	2161	CATGAGCTGCGGAAATATCAGCAAGAAATATTAAGTGAAGTGAAGTGAAGTGAAG	2220
OY	2221	GATTTGTGGAAGCCATGCAATGATCTCTAAAGTGAAGTGAAGTGAAGTGAAGTGAAG	2280
DB	2221	GATTTGTGGAAGCCATGCAATGATCTCTAAAGTGAAGTGAAGTGAAGTGAAGTGAAG	2280
OY	2281	TGAGGGCAGAGAGGCTCTTCTGAGTGTACCACTTGTGCAATTAATGAGGAGGAG	2340
DB	2281	TGAGGGCAGAGAGGCTCTTCTGAGTGTACCACTTGTGCAATTAATGAGGAGGAG	2340
OY	2341	GGAGGGCAGAGAGGCTCTTCTGAGTGTGCAAGTGTGAGGAGTGAAGTGAAGTGAAG	2400
DB	2341	GGAGGGCAGAGAGGCTCTTCTGAGTGTGCAAGTGTGAGGAGTGAAGTGAAGTGAAG	2400
OY	2401	TGACAGAGAGAGTCAAGATTCCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	2460
DB	2401	TGACAGAGAGAGTCAAGATTCCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	2460
OY	2461	GCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2520
DB	2461	GCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2520
OY	2521	CGTATGCTTCAAGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT	2580
DB	2521	CGTATGCTTCAAGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT	2580
OY	2581	AGCAAGTGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	2640
DB	2581	AGCAAGTGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	2640
OY	2641	AAACAGCTTTTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	2700
DB	2641	AAACAGCTTTTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	2700
OY	2701	CGAAGAGTCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2760
DB	2701	CGAAGAGTCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2760
OY	2761	CGGTAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2820
DB	2761	CGGTAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2820
OY	2821	CTTGAAGAGTCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2880
DB	2821	CTTGAAGAGTCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2880
OY	2881	GCCAGTGGCCAGAGTCAACCTCCCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	2940
DB	2881	GCCAGTGGCCAGAGTCAACCTCCCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	2940
OY	2941	TGTGTGGGGCGGAGCAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	3000
DB	2941	TGTGTGGGGCGGAGCAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	3000
OY	3001	CTGGCCCTTGCAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	3060
DB	3001	CTGGCCCTTGCAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	3060
OY	3061	TAGAGTGGCCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3120
DB	3061	TAGAGTGGCCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3120
OY	3121	GTAAGTGAAGTGGTCTTCTTCCCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	3180
DB	3121	GTAAGTGAAGTGGTCTTCTTCCCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	3180
OY	3181	ACCAAAAG	3240
DB	3181	ACCAAAAG	3240
OY	3241	GTAAGTGAAGTGGTCTTCTTCCCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	3300
DB	3241	GTAAGTGAAGTGGTCTTCTTCCCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	3300
OY	3301	CATTGCTGAGAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	3360
DB	3301	CATTGCTGAGAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	3360
OY	3361	AAAAATTCCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3420
DB	3361	AAAAATTCCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3420
OY	3421	TGACAGTCTTCCCTGCTTGTGTAAGAGGCTGAAGTCTGTTGGTCTTGTACATTTCC	3480
DB	3421	TGACAGTCTTCCCTGCTTGTGTAAGAGGCTGAAGTCTGTTGGTCTTGTACATTTCC	3480

THIS PAGE BLANK (USPTO)

Tue Jul 22 16:21:53 2003

us-09-72

Db 3421 TGGCACCTTCCCCTTGCTTTGGTACAAGGGCTGAAGTCCTGTTGGTCTTGTAGCATTTCC 3480
Qy 3481 CAGGATGATGATGTGTCAGCAGGGATGACATCACCACCTTTAGGGCTTTTCCCTGGCAGGGG 3540
Db 3481 CAGGATGATGATGTGTCAGCAGGGATGACATCACCACCTTTAGGGCTTTTCCCTGGCAGGGG 3540
Qy 3541 CCCATGTGGCTAGTCCTCACGAAGACTGGAGTAGAATGTTTGGAGCTCAGGAAGGGTGGG 3600
Db 3541 CCCATGTGGCTAGTCCTCACGAAGACTGGAGTAGAATGTTTGGAGCTCAGGAAGGGTGGG 3600
Qy 3601 TGGAGTGGCCCTCTTCCAGGTGTGAGGGATACGAAGGAGGAAGCTTAGGGAAATCCATTC 3660
Db 3601 TGGAGTGGCCCTCTTCCAGGTGTGAGGGATACGAAGGAGGAAGCTTAGGGAAATCCATTC 3660
Qy 3661 CCCACTCCCTCTTGCCAAATGAGGGGCCAGTCCCCAACAGCTCAGGTCCCCAGAACCCC 3720
Db 3661 CCCACTCCCTCTTGCCAAATGAGGGGCCAGTCCCCAACAGCTCAGGTCCCCAGAACCCC 3720
Qy 3721 CTAGTTCCTCATGAGAAGCTAGGACCAGAAGCACATCGTTCCCTTATCTGAGCAGTGTT 3780
Db 3721 CTAGTTCCTCATGAGAAGCTAGGACCAGAAGCACATCGTTCCCTTATCTGAGCAGTGTT 3780
Qy 3781 TGGGGAACACTACAGTGAACACCTTCTGGAGATGTTAAAGCTTTTACCCACGATAGATT 3840
Db 3781 TGGGGAACACTACAGTGAACACCTTCTGGAGATGTTAAAGCTTTTACCCACGATAGATT 3840
Qy 3841 GTGTTTTAAGGGGTGCTTTTTTAGGGGCATCACTGGAGATAAGAAAGCTGCATTTCAG 3900
Db 3841 GTGTTTTAAGGGGTGCTTTTTTAGGGGCATCACTGGAGATAAGAAAGCTGCATTTCAG 3900
Qy 3901 AAATGCCATCGTAATGGTTTTTAAACACCTTTACCTAATTACAGGTGCTATTTTATAGA 3960
Db 3901 AAATGCCATCGTAATGGTTTTTAAACACCTTTACCTAATTACAGGTGCTATTTTATAGA 3960
Qy 3961 AGCAGACAACACTTCTTTTATGACTCTCAGACTTCTATTTTCATGTTACCATTTTTTTT 4020
Db 3961 AGCAGACAACACTTCTTTTATGACTCTCAGACTTCTATTTTCATGTTACCATTTTTTTT 4020
Qy 4021 GTAACGCAAGGTGTGGGCTTTTGTAACCTCACAGGTGTGGGAGAGACTGCCTTGTTT 4080
Db 4021 GTAACGCAAGGTGTGGGCTTTTGTAACCTCACAGGTGTGGGAGAGACTGCCTTGTTT 4080
Qy 4081 CAACAGTTTGCTCCACTGGTTTCTAATTTTAGGTGCAAAGATGACAGATGCCCAGAGT 4140
Db 4081 CAACAGTTTGCTCCACTGGTTTCTAATTTTAGGTGCAAAGATGACAGATGCCCAGAGT 4140
Qy 4141 TTACCTTTCTGGTTGATTAAAGTTGTATTTCTCTAAAAAAAAAAAAAAAAAAAA 4195
Db 4141 TTACCTTTCTGGTTGATTAAAGTTGTATTTCTCTAAAAAAAAAAAAAAAAAAAA 4195

RESULT 2
AF151969
LOCUS AF151969
SEQUENCE Mus musculus p2h

THIS PAGE BLANK (USPTO)

493 TGGTTGATTAAGTTGATTCTCTCTAATAAAAAAAAAAAAAAAAAAAAAA 4338

RESULT 3
AF068627
LOCUS
DEFINITION Mus musculus DNA cytosine-5 methyltransferase 352 (Dnmt3b) mRNA,
alternatively spliced, complete cds.
AF068627
ACCESSION
AF068627.2 GI:6449471
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Okano, M., Xie, S., and Li, E.
TITLE Cloning and characterization of a family of novel mammalian DNA
JOURNAL Nat. Genet. 19 (3), 219-220 (1998)
MEDLINE
PUBMED 98324766
9662389
REFERENCE
AUTHORS Xie, S., Okano, M., and Li, E.
TITLE Direct Submission
JOURNAL Submitted (28-MAY-1998) CVRC, Mass. Gen. Hospital, 149 13th Street,
Charlestown, MA 02129, USA
3 (bases 1 to 4135)
Okano, M., Chijiwa, T., Sasaki, H., and Li, E.
JOURNAL Submitted (04-NOV-1999) CVRC, Mass. Gen. Hospital, 149 13th Street,
Charlestown, MA 02129, USA
REMARK
COMMENT Sequence update by submitter
On Nov 18, 1999 this sequence version replaced gi:3327981.
FEATURES
source
1.4135
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="2"
/map="84.0 CM"
1..4135
/gene="Dnmt3b"
1..4135
misc_feature
1..4135

THIS PAGE BLANK (USPTO)

/gene="Dm13b"
/note="similar to EST sequences deposited in Genbank
Accession Numbers A11694, A11979, A117277, A120568,
A407106, and A4575617"
269..2788
CDS

/gene="Dm13b"
/function="de novo DNA methylation."
/note="alternatively spliced product: contains Cys-rich
region: C-terminal region is similar to corresponding
region of DNA cytosine-5 methyltransferase 3A"
/codon_start=1

/product="DNA cytosine-5 methyltransferase 3B2"
/protein_id="AAC40179.2"

/db_xref="GI:6449472"

/translation="MKGDSRHILNEBGRASGYECITVNGNFDSDSDYKADSPVLE
ALCEPVCPTETGRSSRLSKREVSLANTODMTGDGDDVDNDGSDILMPK
LRTREKTRRESAPVTRHNSGTSLEORASPTLRGGRHNOYRPEVEPATR
SRRRRSSASPTSPMSAPVDMEETPTSVSPYDLSDGDOEEMDTQVDAESRD
GSTEEODKKEFGIGLVGKIKGFSWNPAMVSWKATSKROAMPMPVONFQDCKP
SRISADKLVALGELFSDHFLATFNKLVSRKAMHRTLEARRARAGCTPSSQESLED
QLKPHLEAHGCFKPEGLGKPNKQEPNKSRRRTNDSASESPPRKLTNSYGG
KORGEDEBSREHMASEVITNNKGNLEDRLCSGKNPVSHPLEEGELCOSCRDREL
FYMYDEGQYOSYCTVCEGRELILCSNCSGRCFCVECLEVLVAGTAADAKIOEWS
CYMGLPORCHGYLRRRKDMNMLDDEFTPDLEEEPEPKLVPAPAAARRPRIVSL
PDGIANTGYLVLELGIKVEKYIASVCASISVGTYKHGGOIKYVNDYKIRKKNIE
WCPDPLVITGSPCNDLSVNPAPRKLEYBETGRLEFEFTHLNTTRKECDNRPTEME
ENVAAMKAVDKDKDISKFLACNPVIMIDALIKVSAHARRVFMGLPGNNRVAMSKNDKL
ELQDLCESTRALKKRVQITTPKINSIRGKNQLEPVPVNGMDVLMCTELERIEFGP
AHYTVSNNGRGAOKILGRSWSVPVIRHLFAPLKDYFCE"

BASE COUNT 1043 a 1048 c 1106 g 938 t
ORIGIN

Query Match 96.9%; Score 4065; DB 10; Length 4135;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 4135; Conservative 0; Mismatches 0; Indels 60; Gaps 1;

QY 1 GAATTCGGGCGCGGGGTTTAAAGCGGCCCAAGTAAAGCGACAGATCGGCGGCGGA 60
DB 1 GAATTCGGGCGCGGGGTTTAAAGCGGCCCAAGTAAAGCGACAGATCGGCGGCGGA 60
QY 61 GATTCGGAACCGGACACTCCGCGCCCGCGCGCGGCGGCGGCGGCGGCGGCGGCGG 120
DB 61 GATTCGGAACCGGACACTCCGCGCCCGCGCGCGGCGGCGGCGGCGGCGGCGGCGG 120
QY 121 GCCGCGCTACAGCCAGCTCAGACAGCGCCCGCTGAGGCTGTGCCAGACCTTGAAC 180
DB 121 GCCGCGCTACAGCCAGCTCAGACAGCGCCCGCTGAGGCTGTGCCAGACCTTGAAC 180
QY 181 TCAGGTATATACCTTCCAGAGCGGGGATCTCCGCCGCCATCCATATGCTTGGGAC 240
DB 181 TCAGGTATATACCTTCCAGAGCGGGGATCTCCGCCGCCATCCATATGCTTGGGAC 240
QY 241 CAATTCAGAGGCTTCTTTCAGAGAAACATGAAGAGAGAGACAGACATCTGAATGA 300
DB 241 CAATTCAGAGGCTTCTTTCAGAGAAACATGAAGAGAGAGACAGACATCTGAATGA 300
QY 301 AGAGGTCGACCGGGTATAGAGAGTGCATATCGTTAATGGAACTTCAAGTACCACT 360
DB 301 AGAGGTCGACCGGGTATAGAGAGTGCATATCGTTAATGGAACTTCAAGTACCACT 360
QY 361 CTCAGACAGAGAGATGCTTCCATCCGCCAGTCTTGGAGGCAATTCGACAGAGCA 420
DB 361 CTCAGACAGAGAGATGCTTCCATCCGCCAGTCTTGGAGGCAATTCGACAGAGCA 420
QY 421 CTCGACACCGAGAGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
DB 421 CTCGACACCGAGAGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
QY 481 CAGCCTTCTGAATTAACAGCAGAGATGACAGAGATGAGAGACAGAGATGATGAAGT 540
DB 481 CAGCCTTCTGAATTAACAGCAGAGATGACAGAGATGAGAGACAGAGATGATGAAGT 540
QY 541 TGATGGAAATGGCTGTGATATCTTAATGCCAAAGCTCACCGCTGAGACCAAGACAC 600

DB 541 TGATGGAAATGGCTGTGATATCTTAATGCCAAAGCTCACCGCTGAGACCAAGACAC 600
QY 601 GACGCGCTCTGAAGAGCGCGGCTCTCCGAACCCGACATAGCATGGAGACCTCCAGCTTGA 660
DB 601 GACGCGCTCTGAAGAGCGCGGCTCTCCGAACCCGACATAGCATGGAGACCTCCAGCTTGA 660
QY 661 GAGGCAAAAGAGCTTCCCGCCAGAAATCAACCCGAGGTGCGGAGGCGGCGGCGGCGG 720
DB 661 GAGGCAAAAGAGCTTCCCGCCAGAAATCAACCCGAGGTGCGGAGGCGGCGGCGGCGG 720
QY 721 GTACCCCTGTGGAGATTTCCGCGCTACAGAGTCTGGAAGACGTGACAGATCGCTTAC 780
DB 721 GTACCCCTGTGGAGATTTCCGCGCTACAGAGTCTGGAAGACGTGACAGATCGCTTAC 780
QY 781 CAGCCATGCTCATCCCTGCGGCTACAGAGTCTGGAAGACGTGACAGATCGCTTAC 840
DB 781 CAGCCATGCTCATCCCTGCGGCTACAGAGTCTGGAAGACGTGACAGATCGCTTAC 840
QY 841 CAGTACCCATCATCTGACTTGAAGCAGATGAGATGAGAGGATATGATACACACA 900
DB 841 CAGTACCCATCATCTGACTTGAAGCAGATGAGATGAGAGGATATGATACACACA 900
QY 901 GGTGATGAGAGAGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 960
DB 901 GGTGATGAGAGAGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 960
QY 961 AATAGTGAACCTGCTGAGGAGAAATCAAGGCTTCTGCTGCTGCTGCTGCTGCTGCT 1020
DB 961 AATAGTGAACCTGCTGAGGAGAAATCAAGGCTTCTGCTGCTGCTGCTGCTGCTGCT 1020
QY 1021 GTCTGGAAGAGCAGCTCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
DB 1021 GTCTGGAAGAGCAGCTCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY 1081 TGCTGATGAG 1140
DB 1081 TGCTGATGAG 1140
QY 1141 CCAGCACTTAAATCTGCTACCTTCAATTAAGCTGCTTCTTATAGAGAGAGAGAGAG 1200
DB 1141 CCAGCACTTAAATCTGCTACCTTCAATTAAGCTGCTTCTTATAGAGAGAGAGAGAG 1200
QY 1201 CACTCTGAG 1260
DB 1201 CACTCTGAG 1260
QY 1261 ACTGAG 1320
DB 1261 ACTGAG 1320
QY 1321 GATCGAG 1380
DB 1321 GATCGAG 1380
QY 1381 TTGAG 1440
DB 1381 TTGAG 1440
QY 1441 CAATGACCTGCTGCTTCTGAGTCCCGCCAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
DB 1441 CAATGACCTGCTGCTTCTGAGTCCCGCCAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
QY 1501 CGGGAAG 1560
DB 1501 CGGGAAG 1560
QY 1561 CAGCAAG 1620
DB 1561 CAGCAAG 1620
QY 1621 CCAGCCCTCTTGTGAGGAGGCTGTGAGAGTTCGCGGAGTCTGCTTCTGAGAGCTCT 1680
DB 1621 CCAGCCCTCTTGTGAGGAGGCTGTGAGAGTTCGCGGAGTCTGCTTCTGAGAGCTCT 1680

THIS PAGE BLANK (USPTO)

```
Db 1561 CCAACCCCTCTTGGAGGGGCTCTGTGAGAGTTGCCGGGATCGCTTCTAGACGCTCTT 1620
Oy 1681 CTACATGTATGATGAGAGCGGCTATGATGCTTACTGACCGTGTGTGTGAGGGCCGTGA 1740
Db 1621 CTACATGTATGATGAGAGCGGCTATGATGCTTACTGACCGTGTGTGTGAGGGCCGTGA 1680
Oy 1741 ACTGCTGCTGTGAGTAAACACAAAGCTGTGAGATGCTTGTGTGTGAGTGTCTGAGGT 1800
Db 1681 ACTGCTGCTGTGAGTAAACACAAAGCTGTGAGATGCTTGTGTGTGAGTGTCTGAGGT 1740
Oy 1801 GCTGTGTGGGGCCAGGACAGCTGTGAGATGCCAACTGTGAGAACCCGTGAGACTGCTATAT 1860
Db 1741 GCTGTGTGGGGCCAGGACAGCTGTGAGATGCCAACTGTGAGAACCCGTGAGACTGCTATAT 1800
Oy 1861 GTGCTGCTCTGAGGCTGTGCTATGGGGTCTCTCGACGCGAGAAAGATTGGAAATGCGGCT 1920
Db 1801 GTGCTGCTCTGAGGCTGTGCTATGGGGTCTCTCGACGCGAGAAAGATTGGAAATGCGGCT 1860
Oy 1921 GCAAGACTTCTTCACTACTGATGATCTGACCTGTGAGAAATTTGAGCCACCAAGTTGTACC 1980
Db 1861 GCAAGACTTCTTCACTACTGATGATCTGACCTGTGAGAAATTTGAGCCACCAAGTTGTACC 1920
Oy 1981 AGCAATTCCTGACAGCAAAAGAGGGCCATTAGAGTCTGTCTGTGTGATGGAATTGC 2040
Db 1921 AGCAATTCCTGACAGCAAAAGAGGGCCATTAGAGTCTGTCTGTGTGATGGAATTGC 1980
Oy 2041 AACGGGCTACTGTGTGCTCAAGAGAGTTGGTATTAAAGTGGAAAGTACATGCTCCGA 2100
Db 1981 AACGGGCTACTGTGTGCTCAAGAGAGTTGGTATTAAAGTGGAAAGTACATGCTCCGA 2040
Oy 2101 AGTGTGTGACAGTCCATGCTGTGTGGAATCTTAAAGTATGAGGCCAGATCAATATGT 2160
Db 2041 AGTGTGTGACAGTCCATGCTGTGTGGAATCTTAAAGTATGAGGCCAGATCAATATGT 2100
Oy 2161 CAATACGTCGGGAAATACCAAGAAATATTTGAAGAGTGGGGCCGTTGACTGTGT 2220
Db 2101 CAATACGTCGGGAAATACCAAGAAATATTTGAAGAGTGGGGCCGTTGACTGTGT 2160
Oy 2221 GATTGGTGGAGCCCATGCAATGATCTCTTAACGTCATCTGCCCGCAAGGTTTATA 2280
Db 2161 GATTGGTGGAGCCCATGCAATGATCTCTTAACGTCATCTGCCCGCAAGGTTTATA 2220
Oy 2281 TGAGGGCAGAGAGAGGCTTCTGTGAGTTTACCACTGTGGAATTAACCGGCCCA 2340
Db 2221 TGAGGGCAGAGAGAGGCTTCTGTGAGTTTACCACTGTGGAATTAACCGGCCCA 2280
Oy 2341 GGGGGCGAAGCCGTCATCTCTGTGAGTGTTCGAGATGTTGGCCATGAAAGTGA 2400
Db 2281 GGGGGCGAAGCCGTCATCTCTGTGAGTGTTCGAGATGTTGGCCATGAAAGTGA 2340
Oy 2401 TGACAGAAAGACATCTCAAGATCTGTGCAATGTAAACCAAGTATGATGATGCCATCA 2460
Db 2341 TGACAGAAAGACATCTCAAGATCTGTGCAATGTAAACCAAGTATGATGATGCCATCA 2400
Oy 2461 GGTGTGTCTGCTCAAGGGCCCGGTACTTCTGGGGTAAACCTACCGGAGTGAAGGCC 2520
Db 2401 GGTGTGTCTGCTCAAGGGCCCGGTACTTCTGGGGTAAACCTACCGGAGTGAAGGCC 2460
Oy 2521 CGTATGAGCTTCAAAAGATGATAGCTCGAGCTGCAAGACTGCGTGAAGTTCAGTAGAC 2580
Db 2461 CGTATGAGCTTCAAAAGATGATAGCTCGAGCTGCAAGACTGCGTGAAGTTCAGTAGAC 2520
Oy 2581 AGCAAAAGTTAAAGAAAGTGCAGCAATTAACCAAGTGCATCCATCAGACAGGCAA 2640
Db 2521 AGCAAAAGTTAAAGAAAGTGCAGCAATTAACCAAGTGCATCCATCAGACAGGCAA 2580
Oy 2641 AAACCGAGCTTCCCTGTACTGATGATGAGTGGCAAGAGAGAGGAGTGTGGGCACTAGCT 2700
Db 2581 AAACCGAGCTTCCCTGTACTGATGATGAGTGGCAAGAGAGGAGTGTGGGCACTAGCT 2640
Oy 2701 CGAAGAGTCTGTGGCTTCCCTGTACTGATGATGAGTGGCAAGAGAGTGTGGGCACTAGCT 2760
Db 2641 CGAAGAGTCTGTGGCTTCCCTGTACTGATGATGAGTGGCAAGAGAGTGTGGGCACTAGCT 2700
```

```
Oy 2761 CCGTCAGAAAGTGTGGGCAAGTCCGTGAGTATACGGGTGATGACAGACCTGTTGGCCC 2820
Db 2701 CCGTCAGAAAGTGTGGGCAAGTCCGTGAGTATACGGGTGATGACAGACCTGTTGGCCC 2760
Oy 2821 CTTGAGAGTACTTTCCTGTGAAATGTTCTACCCAGAGACTGGGAGCTCTCGGTGAGA 2880
Db 2761 CTTGAGAGTACTTTCCTGTGAAATGTTCTACCCAGAGACTGGGAGCTCTCGGTGAGA 2820
Oy 2881 GCCAGTCCCGAGATGACCCCTCCGTGAAGGACCTACCTGCTCCCTTTTAACTAGCTACC 2940
Db 2821 GCCAGTCCCGAGATGACCCCTCCGTGAAGGACCTACCTGCTCCCTTTTAACTAGCTACC 2880
Oy 2941 TGTGTGGGGCTTCACATACAGTACCTGAGTTTCTCTGTGCTAGTGGAGACAGGCTC 3000
Db 2881 TGTGTGGGGCTTCACATACAGTACCTGAGTTTCTCTGTGCTAGTGGAGACAGGCTC 2940
Oy 3001 CTGGCCCTTGCAGAGGAGCCCGGTGCTCCCTGCTGTGACAGCTGAGACTGCTGCT 3060
Db 2941 CTGGCCCTTGCAGAGGAGCCCGGTGCTCCCTGCTGTGACAGCTGAGACTGCTGCT 3000
Oy 3061 TAGAGTAGCCCGCATGCTGCTCATGTTCTTACCTGAAACTTTAAACTTGAAGTAG 3120
Db 3001 TAGAGTAGCCCGCATGCTGCTCATGTTCTTACCTGAAACTTTAAACTTGAAGTAG 3060
Oy 3121 GTAGTAGATGAGCTTCTTTTACCCTCCTGAGTTTATCAGTAAGTATGATGCTAAGAT 3180
Db 3061 GTAGTAGATGAGCTTCTTTTACCCTCCTGAGTTTATCAGTAAGTATGATGCTAAGAT 3120
Oy 3181 ACCAAAAAAGCAAAACAGAAACAGAAACAAAAAACCCTCAACAGCTCTCTTA 3240
Db 3121 ACCAAAAAAGCAAAACAGAAACAGAAACAAAAAACCCTCAACAGCTCTCTTA 3180
Oy 3241 GTACTCAGGTTCATGCTGCAAAATCAGTGGATTTGTTTAAAGTCCGTCTCA 3300
Db 3181 GTACTCAGGTTCATGCTGCAAAATCAGTGGATTTGTTTAAAGTCCGTCTCA 3240
Oy 3301 CATTTGCTGAGAGATGCTATGTAATGTGGGCTCAGATGACAAAGTCAAGGGCCAAA 3360
Db 3241 CATTTGCTGAGAGATGCTATGTAATGTGGGCTCAGATGACAAAGTCAAGGGCCAAA 3300
Oy 3361 AAAAATTCCTCTCCCTCCCGAGAGTATTTAAAGTATGTTTAAAGTCTTCC 3420
Db 3301 AAAAATTCCTCTCCCTCCCGAGAGTATTTAAAGTATGTTTAAAGTCTTCC 3360
Oy 3421 TGGCACCTTCCCTGCTTGTGTAAGAGGCTGGAAGTCTGTGCTGTAGCATTTCC 3480
Db 3361 TGGCACCTTCCCTGCTTGTGTAAGAGGCTGGAAGTCTGTGCTGTAGCATTTCC 3420
Oy 3481 CAGAGTATGATGTCAGAGGATGATCACCACCTTAAAGGCTTTCCCTGCGAGGG 3540
Db 3421 CAGAGTATGATGTCAGAGGATGATCACCACCTTAAAGGCTTTCCCTGCGAGGG 3480
Oy 3541 CCCATGTGCTAGTCTCTCAAGAACTGGAAGTATGTTTGGAGTCAAGAAAGGCTGG 3600
Db 3481 CCCATGTGCTAGTCTCTCAAGAACTGGAAGTATGTTTGGAGTCAAGAAAGGCTGG 3540
Oy 3601 TGGAGTGGGCTTCTCAGAGTGTGAGGATACGAAGAGAGAGCTTAAAGAAATCCATT 3660
Db 3541 TGGAGTGGGCTTCTCAGAGTGTGAGGATACGAAGAGAGAGCTTAAAGAAATCCATT 3600
Oy 3661 CCCACTCCCTCTTCCCAATGAGGGCCCAAGTCCCAACAGCTCAGTCCCGCAAGACCC 3720
Db 3601 CCCACTCCCTCTTCCCAATGAGGGCCCAAGTCCCAACAGCTCAGTCCCGCAAGACCC 3660
Oy 3721 CTAGTCTCTATAGAGCTAGAGACAGAGACATGCTCCCTTATCTGAGCAGTGT 3780
Db 3661 CTAGTCTCTATAGAGCTAGAGACAGAGACATGCTCCCTTATCTGAGCAGTGT 3720
Oy 3781 TGGGAAAGTCTAGAGAAAGCTTCTGAGAGATGTTAAAGCTTTTAAACCCAGATAGAT 3840
Db 3721 TGGGAAAGTCTAGAGAAAGCTTCTGAGAGATGTTAAAGCTTTTAAACCCAGATAGAT 3780
```

THIS PAGE BLANK (USPTO)

Db

2701
2641 CG

22 16:21:53 2003

us-09-720-

OY 3841 GTGTTTTTAAGGGGCTTTTAAAGGGCATCACTGAGATAGAAAGCTGCATTTCAG 3900
|||
Db 3781 GTGTTTTTAAGGGGCTTTTAAAGGGCATCACTGAGATAGAAAGCTGCATTTCAG 3840
OY 3901 AAATGCCATCGTAATGGTTTTTAAACACCTTTTAACTAATTACAGGTGCTATTTATAGA 3960
|||
Db 3841 AAATGCCATCGTAATGGTTTTTAAACACCTTTTAACTAATTACAGGTGCTATTTATAGA 3900
OY 3961 AGCAGACAACTTCCTTTTAACTCACTCAGACTTCATTTTCATGTTACCATTTT 4020
|||
Db 3901 AGCAGACAACTTCCTTTTAACTCACTCAGACTTCATTTTCATGTTACCATTTT 3960
OY 4021 GTAACGCAAGGTGTGGCTTTTGTAACTCACTCAGAGTGTGGGAGAGACTGCCCTGTT 4080
|||||
Db 3961 GTAACGCAAGGTGTGGCTTTTGTAACTCACTCAGAGTGTGGGAGAGACTGCCCTGTT 4020
OY 4081 CAACAGTTGTCTCCACTGTTCTAATTTTAAAGTCAAGATGACAGATGCCAGAGT 4140
|||
Db 4021 CAACAGTTGTCTCCACTGTTCTAATTTTAAAGTCAAGATGACAGATGCCAGAGT 4080
OY 4141 TTACCTTCTGTTGATTAAAGTGTATTCTCTAATAAAAAAAAAAAAAA 4195
|||||
Db 4081 TTACCTTCTGTTGATTAAAGTGTATTCTCTAATAAAAAAAAAAAAAA 4135

THIS PAGE BLANK (USPTO)



score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

OM nucleic - nucleic search, using sw model

17408.415 Million cell updates/sec

PERFECT SCORE: 4255

Sequence: 1 gccgcgcaaccaggcgcc.....taaaaaaaaaaaaaaa 4293

Scoring table:

IDENTITY_NUC	Correct	100
0	1	0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

GenEmbl: *

- 1: gb_ba: *
- 2: gb_hig: *
- 3: gb_in: *
- 4: gb_om: *
- 5: gb_ov: *
- 6: gb_pat: *
- 7: gb_ph: *
- 8: gb_pl: *
- 9: gb_pri: *
- 10: gb_pro: *
- 11: gb_sus: *
- 12: gb_sy: *
- 13: gb_un: *
- 14: gb_vl: *
- 15: em_ba: *
- 16: em_fun: *
- 17: em_hum: *
- 18: em_in: *
- 19: em_mu: *
- 20: em_om: *
- 21: em_or: *
- 22: em_ov: *
- 23: em_pat: *
- 24: em_ph: *
- 25: em_pl: *
- 26: em_pro: *
- 27: em_sus: *
- 28: em_un: *
- 29: em_vl: *
- 30: em_hig_hum: *
- 31: em_hig_in: *
- 32: em_hig_other: *
- 33: em_hig_mus: *
- 34: em_hig_pln: *
- 35: em_hig_rod: *
- 36: em_hig_mam: *
- 37: em_hig_vrt: *
- 38: em_sy: *
- 39: em_hgo_hum: *
- 40: em_hgo_mus: *
- 41: em_hgo_other: *

Pred. No. is the number of results predicted by chance to have a

Result	Score	Query	ID	Description
No.		Match Length DB		
C 1	4180.6	97.4	4258	9 AF331856 Homo sapi
2	3005	70.0	3005	9 AF067972 Homo sapi
3	2879.2	65.1	4192	10 AF068625 Mus muscu
4	2825.4	67.8	4094	10 BC007466 Homo sapi
5	2139.6	49.8	2191	9 AK025230 Homo sapi
6	1667	38.8	1756	9 BC018214 Homo sapi
7	1408.6	32.8	137652	9 AC012074 Homo sapi
8	1221.4	28.5	123936	2 AC120624 Rattus no
9	1157.6	27.0	168651	9 AC009474 Homo sapi
C 10	1078.4	25.1	176697	2 AC112586 Rattus no
11	791.8	18.4	4267	9 AF176228 Homo sapi
12	788.2	18.4	4135	10 AF068627 Mus muscu
13	788.2	18.4	4163	10 AF151974 Mus muscu
14	786.6	18.3	4278	10 AF151970 Mus muscu
15	764.4	17.8	4145	9 AF156488 Homo sapi
16	764.4	17.8	4335	9 AF331857 Homo sapi
17	762	17.7	4195	10 AF068626 Mus muscu
18	760.4	17.7	4223	10 AF151973 Mus muscu
19	760.4	17.7	4338	10 AF151969 Mus muscu
C 20	694	16.2	119630	2 AC112040 Rattus no
C 21	671.6	15.6	117264	2 AC103454 Rattus no
22	647	15.1	204534	2 AC118195 Mus muscu
23	647	15.1	208874	2 AC111092 Mus muscu
24	601.4	14.0	3897	9 AF156487 Mus muscu
25	600	14.0	3946	10 AF068628 Homo sapi
26	600	14.0	3974	10 AF151976 Mus muscu
27	598.4	13.9	4089	10 AF151971 Mus muscu
28	573.8	13.4	4006	10 AY078427 Mus muscu
29	572.2	13.3	4034	10 AF151975 Mus muscu
30	572.2	13.3	4149	10 AF151972 Mus muscu
31	544	12.7	2848	5 AF135438 Dantio rer
C 32	533.2	12.4	210269	5 AF135459 Mus muscu
33	492.6	11.5	119630	2 AC112040 Rattus no
34	492.2	11.5	3017	9 AK001191 Homo sapi
35	417	9.7	167566	2 AC111734 Rattus no
36	411.2	9.6	2127	9 AF129267 Homo sapi
37	374.6	7.6	2077	6 AR129189 Sequence
38	294.8	6.9	2057	9 AF129268 Homo sapi
C 39	273	6.4	273	6 AA332056 Sequence
C 40	273	6.4	273	6 AX431206 Sequence
C 41	273	6.4	273	6 GG6200 human STR
C 42	254.8	5.9	56798	2 AC034293 Homo sapi
C 43	252	5.9	2008	9 AF129269 Homo sapi
C 44	190.6	4.4	204534	9 AC118195 Mus muscu
C 45	190.6	4.4	208874	2 AC111092 Mus muscu

RESULT 1

LOCUS	AF331856	4258 bp	mRNA	linear	PRI 02-JAN-2007
DEFINITION	Homo sapiens DNA cytosine methyltransferase 3 alpha (DNMT3A) mRNA,				

LABORATORY

VERSION AF331856.1 GI:18033252

SOURCE

ORGANIS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 4258)

TITLE Cloning, expression and characterization of human DNMT3 genes

JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 4258)
 AUTHORS M.J., Pradhan, S. and Roberts, R.J.
 TITLE Direct Submission
 JOURNAL Submitted (22-DEC-2000) New England Biolabs, 32 Tozer Road,
 Beverly, MA 01915, USA
 FEATURES
 source Location/Qualifiers
 1..4258
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="2"
 1..4258
 /gene="DNMT3A"
 227..2956
 /gene="DNMT3A"
 /codon_start=1
 /product="DNA cytosine methyltransferase 3 alpha"
 /protein_id="AA157039.1"
 /translation="MPSGGPDYSSAAEREDRKDGEDEPRKREPERESTARK
 VRRPGRKRHPPEVSGDTPKDPAYISKSPSMAODSGASSELLPNDLEKRSBPBEGS
 PAGQCGAPAEIGESAEETLPASRAVENGGCTPKGEGAPAEAGKQKETSMM
 EGRRLRGIGMESLRQRPRLTFQAGDPTYSKRKDEMTLRMRKRAEKAYI
 AGMAVEBNOGPESEKVEASPAVOQPTDPASTVATPEVPSADGKNNTKAD
 DEPEYEDRGFGIGELVWGLRGFSWMPGRIVSMNTRSRRAAGTMMVWPFQKES
 YVCEKLPLPSFCSAFHQTNYKQPMYKAIYEVLQVSRACKLEPVCHEDESDT
 AAAYVONKPMIEWALGQFQSGPKLEPEREENKPYEYTDVMVEEAAVAAPPP
 AKRPRSTAEKPKYELIDERTRELVYERQKCNIDICISGSLNVLLEPLVGP
 GACQNCNCFLECAVYDDDOYSCYICCGREVLMQGNMNCRCRCFVLLVGP
 GAAQAIKEDPMNCMGCKGYGLRLREDEPSRLQFEPANNIDFDPKVPYPPV
 AKRKPILVLSLFDGATGLVLDLQIVDRYIASEVCEDSITGVGRHQKIMYV
 DVRSYQKHIOEMFQPDVLYGSDPCNDLSIVNPAKEDSITGVGRHQKIMYV
 KEGDRPFPFLFENVVAMGVSDKRDLSRFLSRLSPMIDAKSVSAHARFYMNLDM
 NPLASTVNDKLELQECLEHGRJAKFSKVRITITRSNSIKQKQHPFVPMNEKEDIL
 NCTEMERFVGPPVHYTDVSNMSRLRQLRSGMSVPVIRHLFAPLEKPYACV"

BASE COUNT 1062 a 1089 c 1257 g 850 t

Query Match 97.4% Score 4180.6; DB 9; Length 4258;
 Best Local Similarity 99.7% Pred. No. 0;
 Matches 4210; Conservative 0; Mismatches 9; Indels 4; Gaps 2;

OY 52 GGTGAGCCATGGAAGCCCCACCCACAGGCTGACAGAGCCCTTACAGAGGCTC 111
 DB 40 GGTGAGCCATGGAAGCCCCACCCACAGGCTGACAGAGCCCTTACAGAGGCTC 99
 OY 112 AACACGGGATCTATGTTTAAGTTTAACTCTGCTCCCAAGACCAAGATATTCCTTC 171
 DB 100 AACACGGGATCTATGTTTAAGTTTAACTCTGCTCCCAAGACCAAGATATTCCTTC 159
 OY 172 CCCAAAGCCAGAGCCCCCGAGCCCGCGACAGCCCAAGCTGCTCCCGCGCCAGAT 231
 DB 160 CCCAAAGCCAGAGCCCCCGAGCCCGCGACAGCCCAAGCTGCTCCCGCGCCAGAT 219
 OY 232 GCCCGCATGCTCTCAAGCGGCCCGGGGACACCAAGCTGCTGCGAGCGGAGGA 291
 DB 220 GCCCGCATGCTCTCAAGCGGCCCGGGGACACCAAGCTGCTGCGAGCGGAGGA 279
 OY 292 GGACGGAAAGAGCAGAGAGAGAGAGAGCCGCTGGCAAGAGAGAGCGCCCAAGGCC 351
 DB 280 GGACGGAAAGAGCAGAGAGAGAGAGAGCCGCTGGCAAGAGAGAGCGCCCAAGGCC 339
 OY 352 CAGACACAGCGAGCGAAGGTGGGGCGGCTGGAGGAAGCCCAAGCACCCCGGCTGA 411
 DB 340 CAGACACAGCGAGCGAAGGTGGGGCGGCTGGAGGAAGCCCAAGCACCCCGGCTGA 399
 OY 412 AAGCGGTGACAGCGCCAAAGACCTCGGGTGATCTTCCAAAGTCCCATTCATGGCCAGGA 471
 DB 400 AAGCGGTGACAGCGCCAAAGACCTCGGGTGATCTTCCAAAGTCCCATTCATGGCCAGGA 459
 OY 472 CTCAGAGCGCTCAGAGCTATTACCAATGGAGGACTTGGAGGAAGCGAGTACGCCAGCC 531
 DB 460 CTCAGAGCGCTCAGAGCTATTACCAATGGAGGACTTGGAGGAAGCGAGTACGCCAGCC 519

OY 532 ACAGAGAGGAGAGCCCTGCTGGGGGGGACGAAGAGCGGGGCCCCACAGAGAGGAGGTGC 591
 DB 520 AGAGAGAGGAGAGCCCTGCTGGGGGGGACGAAGAGCGGGGCCCCACAGAGAGGAGGTGC 579
 OY 592 ACCTGAGAGCCCTGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 651
 DB 580 ACCTGAGAGCCCTGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 639
 OY 652 GGGCGGAGAGAGCCCTGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 711
 DB 640 GGGCGGAGAGAGCCCTGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 699
 OY 712 GAAATGAGAGAGGCTGCTGGGGGGGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 771
 DB 700 GAAATGAGAGAGGCTGCTGGGGGGGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 759
 OY 772 TCAGAGAGGCTGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 831
 DB 760 TCAGAGAGGCTGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 819
 OY 832 CAAGGGGAGAGAGGCTGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 891
 DB 820 CAAGGGGAGAGAGGCTGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 879
 OY 892 TCAGAGAGGCTGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 951
 DB 880 TCAGAGAGGCTGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 939
 OY 952 GGGCGAGGCTGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1011
 DB 940 GGGCGAGGCTGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 999
 OY 1012 GCGTGAAGGCTGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1071
 DB 1000 GCGTGAAGGCTGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1059
 OY 1072 GCGAGAGGCTGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1131
 DB 1060 GCGAGAGGCTGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1119
 OY 1132 GGGCTTCTCTGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1191
 DB 1120 GGGCTTCTCTGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1179
 OY 1192 AGCTGAAGGCTGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1251
 DB 1180 AGCTGAAGGCTGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1239
 OY 1252 TGAGAGGCTGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1311
 DB 1240 TGAGAGGCTGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1299
 OY 1312 GCAGAGGCTGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1371
 DB 1300 GCAGAGGCTGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1359
 OY 1372 GAAGCTTCTCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1431
 DB 1360 GAAGCTTCTCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1419
 OY 1432 GCAGAGGCTGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1491
 DB 1420 GCAGAGGCTGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1479
 OY 1492 CCTGAGAGGCTGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1551
 DB 1480 CCTGAGAGGCTGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1539
 OY 1552 GGAACCTGAGAGGCTGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 1611
 DB 1540 GGAACCTGAGAGGCTGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 1599

OY	1612	AGCGGAGAAAGCCCAAGGTC	CAAGAGATTTATGTAGACCGCA	CAAGAGAGCGGCTGGGTGA	1611
Db	1600	AGCGGAGAAAGCCCAAGGTC	CAAGAGATTTATGTAGACCGCA	CAAGAGAGCGGCTGGGTGA	1609
OY	1672	CGAGGTGGCGCAGAAAGTGC	CGGAACATTTGAGAGCATGTGCAT	CTCCTGTGGAGACCTCCAA	1671
Db	1660	CGAGGTGGCGCAGAAAGTGC	CGGAACATTTGAGAGCATGTGCAT	CTCCTGTGGAGACCTCCAA	1659
OY	1732	TGTTAACCTTGGAAACCCCT	CTTTCGTTTGAGAGATGTGCC	AAATGCAAGAACTGCTT	1731
Db	1720	TGTTAACCTTGGAAACCCCT	CTTTCGTTTGAGAGATGTGCC	AAATGCAAGAACTGCTT	1719
OY	1792	TCTTGAGATGTGCGCTACCA	GTACAGACGACGGCTACCAC	TCTCTACTGTCACACATCTGCTG	1791
Db	1780	TCTTGAGATGTGCGCTACCA	GTACAGACGACGGCTACCAC	TCTCTACTGTCACACATCTGCTG	1779
OY	1852	TGGGGGCGCGTGAAGGTGCT	CATGTCCGGAAACAAAC	TGGTGGCGAGGTCCTTTCG	1851
Db	1840	TGGGGGCGCGTGAAGGTGCT	CATGTCCGGAAACAAAC	TGGTGGCGAGGTCCTTTCG	1839
OY	1912	GTGTGTGAGACCTTGTGTG	TGGGGGCCGCCAGGACGCA	GCATTAAGAGAAAGCCCTTG	1911
Db	1900	GTGTGTGAGACCTTGTGTG	TGGGGGCCGCCAGGACGCA	GCATTAAGAGAAAGCCCTTG	1899
OY	1972	GAACTGCTACATGTGCGGG	CAACAAGGTAACCTACGGG	CTGCTGCGCGCGCAGAGACTG	1971
Db	1960	GAACTGCTACATGTGCGGG	CAACAAGGTAACCTACGGG	CTGCTGCGCGCGCAGAGACTG	1959
OY	2032	GCCCTCCCGGCTCCAGANT	GCTTCGCTTAATTAACAG	ACAGGAATTTGACACCTCC	2031
Db	2020	GCCCTCCCGGCTCCAGANT	GCTTCGCTTAATTAACAG	ACAGGAATTTGACACCTCC	2019
OY	2092	GTTTTTACCACCTTCCAGC	TGTGAGAGAGAAACCCAT	CCGGGTGCTGTCTCTTTGA	2091
Db	2080	GTTTTTACCACCTTCCAGC	TGTGAGAGAGAAACCCAT	CCGGGTGCTGTCTCTTTGA	2079
OY	2152	TGGAATGCTACAGAGGCT	CCTGTGTGTGAAGAGACT	TGGCAATTCAGGTGAGACGCTACAT	2151
Db	2140	TGGAATGCTACAGAGGCT	CCTGTGTGTGAAGAGACT	TGGCAATTCAGGTGAGACGCTACAT	2139
OY	2212	TGCGTCGAGAGTGTGTAG	AGACTCCATTCACGGTGGG	CAATGGTGGCGACCAAGGGAAGAT	2211
Db	2200	TGCGTCGAGAGTGTGTAG	AGACTCCATTCACGGTGGG	CAATGGTGGCGACCAAGGGAAGAT	2199
OY	2272	CATGTACGTCGGGAGCGT	CCGACGCTACACAAAGCAT	ATTCAGAGAGTGGGGCCCAT	2271
Db	2260	CATGTACGTCGGGAGCGT	CCGACGCTACACAAAGCAT	ATTCAGAGAGTGGGGCCCAT	2259
OY	2332	CGATCTGGTGAATTTGGG	GCGACGCTCCGATGACCT	CTCATCTGTCACACCCCTGCTGC	2331
Db	2320	CGATCTGGTGAATTTGGG	GCGACGCTCCGATGACCT	CTCATCTGTCACACCCCTGCTGC	2319
OY	2392	GGGGCTCTACAGAGGGCA	CTGGCGGGCTTTTGTGAG	TTCTACACCGCTCTCTGCAATATGC	2391
Db	2380	GGGGCTCTACAGAGGGCA	CTGGCGGGCTTTTGTGAG	TTCTACACCGCTCTCTGCAATATGC	2379
OY	2452	GGGGCCCAAGAGGAGATG	ATCCGCCCTTTTGGCTCT	TGAGATGAGAGTGGGGCCAT	2451
Db	2440	GGGGCCCAAGAGGAGATG	ATCCGCCCTTTTGGCTCT	TGAGATGAGAGTGGGGCCAT	2439
OY	2512	GGGGCTTGTGACAAAGAG	GACATCTCGCGATTTCTG	AGTCCAAACCCCTGTGATATGA	2511
Db	2500	GGGGCTTGTGACAAAGAG	GACATCTCGCGATTTCTG	AGTCCAAACCCCTGTGATATGA	2499
OY	2572	TGCGCAAAAGAAATGTG	CAGTGCACACAGGCGCC	CTACTTCTGGGGGTAACCTTCCCGGTAT	2571
Db	2560	TGCGCAAAAGAAATGTG	CAGTGCACACAGGCGCC	CTACTTCTGGGGGTAACCTTCCCGGTAT	2559
OY	2632	GAAACAGGCGGTTGGCAT	CTCAGTGTGAAATGAT	TAACTGAGAGCTGCAGAGAGTGTGAGACA	2631
Db	2620	GAAACAGGCGGTTGGCAT	CTCAGTGTGAAATGAT	TAACTGAGAGCTGCAGAGAGTGTGAGACA	2619
OY	2692	TGGCAGGATTAACCAAG	TTCACAAAGTAGAGAC	CAATTAACAGAGTCCAAATCTCATATAA	2691

Db	2680	TTGGCAGGATAGCAAATTACGACAAAGTAGGACCTTATCTACGAGGTCAAACCTCCATAAA	2739
QY	2752	GCAGGGCAAAACACGACATTTTTCCTGTCCTTATGATATGAGAAAGACATCTTATGGTG	2811
Db	2740	GCAGGGCAAAACACGACATTTTTCCTGTCCTTATGATATGAGAAAGACATCTTATGGTG	2799
QY	2812	CACCTGAATATGAAAAGGATATTTGGTTTCCAGTCACATTAATCTGACGCTTCCAAACATGAG	2871
Db	2800	CACCTGAATATGAAAAGGATATTTGGTTTCCAGTCACATTAATCTGACGCTTCCAAACATGAG	2859
QY	2872	CCGCTTTGGCGAGCGAGACACTGCTGGGCGGCTCATGAGCGCTGCCAGTCACTCCGCACT	2931
Db	2860	CCGCTTTGGCGAGCGAGACACTGCTGGGCGGCTCATGAGCGCTGCCAGTCACTCCGCACT	2919
QY	2932	CTTGCGCTCCGCGTGAAGAGATATTTTGGCTGTGTGAAGGACATGAGGGGCAACTGTAGGT	2991
Db	2920	CTTGCGCTCCGCGTGAAGAGATATTTTGGCTGTGTGAAGGACATGAGGGGCAACTGTAGGT	2979
QY	2992	AGCGACACAAAGTTAAACAAACAAACAAACAAACAAACATTAATAAACACGAAGA	3051
Db	2980	AGCGACACAAAGTTAAACAAACAAACAAACAAACAAACATTAATAAACACGAAGA	3039
QY	3052	TGAGATGAGAGAAATTCACGACCCAGAGAGAAAAAGCAATTTAAACAAAACCCAC	3111
Db	3040	TGAGATGAGAGAAATTCACGACCCAGAGAGAAAAAGCAATTTAAACAAAACCCAC	3099
QY	3112	AGAGCGGAAATTAACCGGAGGGCTTGGCTTCCGAAAAGGGTTGGACATCATCTCTGATT	3171
Db	3100	AGAGCGGAAATTAACCGGAGGGCTTGGCTTCCGAAAAGGGTTGGACATCATCTCTGATT	3159
QY	3172	TTTCATGTATATCTTCAGTCCATATTTAAACAAACAAACCAAGCTCCCTCCCTCCCTCC	3231
Db	3160	TTTCATGTATATCTTCAGTCCATATTTAAACAAACAAACCAAGCTCCCTCCCTCCCTCC	3219
QY	3232	CCCTCCCTTTTTCGCTCAGACCTTTATTTCTACTCTTTCAGAGGGGTTTCTGT	3291
Db	3220	CCCTCCCTTTTTCGCTCAGACCTTTATTTCTACTCTTTCAGAGGGGTTTCTGT	3279
QY	3292	TTGTTGGGTTTCTGTTCTGCTGACATGAAACAAGAGGTTATTCAGCAAAAATCAG	3351
Db	3280	TTGTTGGGTTTCTGTTCTGCTGACATGAAACAAGAGGTTATTCAGCAAAAATCAG	3339
QY	3352	TAACAAAATATAGTAAACAATACCTTGACAGAGAAAGTGGGAGAGAGAAAGGGA	3411
Db	3340	TAACAAAATATAGTAAACAATACCTTGACAGAGAAAGTGGGAGAGAGAAAGGGA	3396
QY	3412	ATTTTAAAGAAATCTATATATGGGGTGTATTTTGTATTTTGTATTTTGTATTTTGT	3471
Db	3397	AAATCTATAGAAATCTATATATGGGGTGTATTTTGTATTTTGTATTTTGTATTTTGT	3456
QY	3472	GGGTTTTTTTTTTTACTATATATCTTTTTTGTGTCTCTAGCTGATAGATAGAG	3531
Db	3457	GGGTTTTTTTTT-TTTTACTATATATCTTTTTTGTGTCTCTAGCTGATAGATAGAG	3515
QY	3532	CACAAGCAGGGGAGGAAAGAGAGACACTCAGGCGGACACTTCCCTCCACCACTG	3591
Db	3516	CACAAGCAGGGGAGGAAAGAGAGACACTCAGGCGGACACTTCCCTCCACCACTG	3575
QY	3592	AGCTGTGCTGCCAGCACCATTCTGTGTACGCAAAAACGAACCCAGTTAGCAGAGGAG	3651
Db	3576	AGCTGTGCTGCCAGCACCATTCTGTGTACGCAAAAACGAACCCAGTTAGCAGAGGAG	3635
QY	3652	ACGAAACACCAACAAAGACATTTTTCACAGTATTCAGTGGCTTCCACACGGAAC	3711
Db	3636	ACGAAACACCAACAAAGACATTTTTCACAGTATTCAGTGGCTTCCACACGGAAC	3695
QY	3712	CTTGAGAGAAATCAGTTTCTAGAGCCGCTGTACCTCTGTTACAGTTATATATATA	3771
Db	3696	CTTGAGAGAAATCAGTTTCTAGAGCCGCTGTACCTCTGTTACAGTTATATATATA	3755
QY	3772	TGATATATGAGATATATATATATAAGGTACTGTTAACTACTGACACCCGACTTCAT	3831

[illegible]

gene	1. .3005
CDS	/gene="DNMT3A" 230. .2968 /gene="DNMT3A" /function="de novo DNA cytosine methyltransferase" /codon_start=1 /product="DNA cytosine methyltransferase 3 alpha" /protein_id="AAD33084.2" /db_xref="gi:12746532" /translation="MPAPSSGSGDTSYSAAREDRKDGEDEPRGAEKREOBSTT" ARVGGGGRKHPVESGDTPKDPDAVIVSKRSMAQSGASELIPNDLEKSEBQAP ESGAPGCGGAPAGEEGEAETLPKASRAVENGGCTPEGSGAPAEAGDEKQETIE MKHGESNRLRGGLJMESLRLRQMPRLPFOAGDDPYIVSKRRLKMAEKARKK KVIAGNANLEENOGGESOKYEASPAVQQTDPDASPTVATPEVSDADCKNATY AGDVEYEDDGRGFGTIGLVNGLRQFSAHYKQMPYKALIEYLQVYASRAAGKLPVCHSDSD KSFQVEVKLMLSPSCAFHOATYKQMPYKALIEYLQVYASRAAGKLPVCHSDSD SDPAKREYONKPMLEALGDFGFGSPGPKLEPEERKPIKEAYETVDMVPEALAAVY SPAPKREYONKPMLEALGDFGFGSPGPKLEPEERKPIKEAYETVDMVPEALAAVY FVGGAAONKPNFLECAOYDDGYSOYCTICGSGREVLKMGNNCCDFCEVCDL FVGGAAONKPNFLECAOYDDGYSOYCTICGSGREVLKMGNNCCDFCEVCDL YVDAEKRPIRYLSFDG IATGLVLYKDG I QVDY I IVEYCEDS IYVGGRHOCK I YVDAEKRPIRYLSFDG IATGLVLYKDG I QVDY I IVEYCEDS IYVGGRHOCK I APRKEDDPRFPMLEFENYVAGVSDRKRLSRLEENPMIDAKESVAHRAHPFENMEXI PCNRPLASTVNDKLELOELGHRILAKSKRITITTSNISKQGRHOPFENMEXI DILMCTEMRNVGFPVHYTDSVMSHLAQRLIGLSMWSVPVIRHLPAPKEYFACY"
BASE COUNT	674 a 846 c 978 g 507 t
ORIGIN	
Query Match	70.0%; Score 3005; Db 9; Length 3005;
Best Local Similarity	100.0%; P-Exp. No. 0; Mismatches 0; Gaps 0
Matches 3005; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
QY	1 GCCGGGACACCGAGCGCGACCCCGCGGCCGACCCACCGGCGATACGCTGAGGCC 60
Db	1 GCCGGGACACCGAGCGCGCGACCCCGCGGCCGACCCACCGGCGATACGCTGAGGCC 60
QY	61 ATCGAAGCCCCCACCACCGAGGCTGACAGAGGACACCGTTACCAAGAGGCTCAACCCGG 120
Db	61 ATCGAAGCCCCCACCACCGAGGCTGACAGAGGACACCGTTACCAAGAGGCTCAACCCGG 120
QY	121 ATCTATGTTTAACCTTTTAACCTCTCCCTCCCAAGACACGATTAATTCCTTCCCAAGCC 180
Db	121 ATCTATGTTTAACCTTTTAACCTCTCCCTCCCAAGACACGATTAATTCCTTCCCAAGCC 180
QY	181 CAGCAGCCCCCAGCCCCCGCGAGCGCCAGGCTCTCCGCGCGCCCGAGATGCCGCGCAT 240
Db	181 CAGCAGCCCCCAGCCCCCGCGAGCGCTCTCCGCGCGCCCGAGATGCCGCGCAT 240
QY	241 GCCCTCCAGCGCGCGCGCGGAGCAGCAGAGCTCTGCTCGGAGAGCGGAGAGGACCAAA 300
Db	241 GCCCTCCAGCGCGCGCGGAGCAGCAGAGCTCTGCTCGGAGAGCGGAGAGGACCAAA 300
QY	301 GGCAGAGAGAGAGAGAGAGAGAGAGCGCGTGGCAGAGAGAGCGCAAGAGGCCACACAC 360
Db	301 GGCAGAGAGAGAGAGAGAGAGAGAGCGCGTGGCAGAGAGAGCGCAAGAGGCCACACAC 360
QY	361 GGCACGGAAGTGGGGCGGCTCGGAGAGAACCGCAACGACACCCCGGCTGGAAACGGTGA 420
Db	361 GGCACGGAAGTGGGGCGGCTCGGAGAGAACCGCAACGACACCCCGGCTGGAAACGGTGA 420
QY	421 CACGGCAAAAGACCTTCGCGTGATCTCCAAAGTCCCATTCATTCATTCGCGCAAGACTCAGGCGC 480
Db	421 CACGGCAAAAGACCTTCGCGTGATCTCCAAAGTCCCATTCATTCATTCGCGCAAGACTCAGGCGC 480
QY	481 CTCAGAGCTATTACCAAAATGGGACTTTGGAGAAAGCGAGTGAAGCCCCACCAAGAGAGG 540
Db	481 CTCAGAGCTATTACCAAAATGGGACTTTGGAGAAAGCGAGTGAAGCCCCACCAAGAGAGG 540
QY	541 GAGCCCTCTCTGGGGGAGCAGAGGAGCGGGCCCGCAGCAGAGAGAGAGGCTGCACGCTGAGAC 600
Db	541 GAGCCCTCTCTGGGGGAGCAGAGGAGCGGGCCCGCAGCAGAGAGAGAGGCTGCACGCTGAGAC 600
QY	601 CCGGCTGAAGCCTCAAGAGCAGAGTGGAAAAATGGCTGCTGCACCCCAAGAGAGAGGCGCAGG 660

Db 601 CCTGCTGTAAGCCTCAGAGCAGTGGAAATGGCTGTGCAACCCCAAGAGGGCGGAGG 660
QY 661 AGCCCTGTCAGAGGGGGCAAGAAACAGAGAGAGCCAAATCCGATTCATGAAATGGA 720
Db 661 AGCCCTGTCAGAGGGGGCAAGAAACAGAGAGAGCCAAATCCGATTCATGAAATGGA 720
QY 721 GGGGTCCCGGGGGCGGCTGGGGGGGCTGGGGCTGGAGATCCAGCTCCGTCAGCGGC 780
Db 721 GGGGTCCCGGGGGCGGCTGGGGGGGCTGGGGCTGGAGATCCAGCTCCGTCAGCGGC 780
QY 781 CATGCCAGGCTCACCTTCCAGCGGGGACCCCTACTACATCAGCAAGCCGACGGGA 840
Db 781 CATGCCAGGCTCACCTTCCAGCGGGGACCCCTACTACATCAGCAAGCCGACGGGA 840
QY 841 CGAGTGGCTGGCAGCTGGGAAAGAGGAGCTGAGAAAGCAAGGTCATTGCAAGAT 900
Db 841 CGAGTGGCTGGCAGCTGGGAAAGAGGAGCTGAGAAAGCAAGGTCATTGCAAGAT 900
QY 901 GAATGCTGTGGAAGAAACAGAGGGCCCGGGAGTCTCAGAAAGTGGAGAGGCCAGGCC 960
Db 901 GAATGCTGTGGAAGAAACAGAGGGCCCGGGAGTCTCAGAAAGTGGAGAGGCCAGGCC 960
QY 961 TCCTGCTGTGACAGAGCCCACTGACCCCGCATCCCCCACTGTGGCTACACGCTGAGCC 1020
Db 961 TCCTGCTGTGACAGAGCCCACTGACCCCGCATCCCCCACTGTGGCTACACGCTGAGCC 1020
QY 1021 CGTGGGGTCCGATGCTGGGGCAAGAAATGCCAACCAAGCCATGACAGGCCGAGTA 1080
Db 1021 CGTGGGGTCCGATGCTGGGGCAAGAAATGCCAACCAAGCCGATGACAGGCCGAGTA 1080
QY 1081 CGAGAGCGCCCGGGGCTTTGGCATTTGGGAGCTGGTGGGGAATCGGGGGCTTCTC 1140
Db 1081 CGAGAGCGCCCGGGGCTTTGGCATTTGGGAGCTGGTGGGGAATCGGGGGCTTCTC 1140
QY 1141 CTGGTGGCCAGGCGGCATTTGTCTTGTGGATGACGGGGCCGAGCCGACAGCTGAAG 1200
Db 1141 CTGGTGGCCAGGCGGCATTTGTCTTGTGGATGACGGGGCCGAGCCGACAGCTGAAG 1200
QY 1201 CACCGGCTGGGTCATGTGTTGCGAGAGGCGAAATTCACAGTGGTGTGTTGAAACCT 1260
Db 1201 CACCGGCTGGGTCATGTGTTGCGAGAGGCGAAATTCACAGTGGTGTGTTGAAACCT 1260
QY 1261 GATGCGGCTGAGCTGTTTTCAGTGCCTTCACAGGCGCAGTACAAAGACGCCAT 1320
Db 1261 GATGCGGCTGAGCTGTTTTCAGTGCCTTCACAGGCGCAGTACAAAGACGCCAT 1320
QY 1321 GTACCGCAAGCCATCTACGAGGTCTGCAAGTGGCAGACGCGCGGGGAAGCTGT 1380
Db 1321 GTACCGCAAGCCATCTACGAGGTCTGCAAGTGGCAGACGCGCGGGGAAGCTGT 1380
QY 1381 CCGGGTGGCCACAGACGATGAGAGTGAACCTGCCAAAGCCGCGAGAGTGCAGAAACAA 1440
Db 1381 CCGGGTGGCCACAGACGATGAGAGTGAACCTGCCAAAGCCGCGAGAGTGCAGAAACAA 1440
QY 1441 GCCCATGATTAATGGGCGCTGGGGGCTTCAGGCTTCTGGCCCTTAAGGGCTGAGGC 1500
Db 1441 GCCCATGATTAATGGGCGCTGGGGGCTTCAGGCTTCTGGCCCTTAAGGGCTGAGGC 1500
QY 1501 ACCAGAGAGAGAGAAATCCCTACAAAGAGTGTACACGAGCATGTGGTGAACCTGA 1560
Db 1501 ACCAGAGAGAGAGAAATCCCTACAAAGAGTGTACACGAGCATGTGGTGAACCTGA 1560
QY 1561 GGGAGCTGCTACGACACCACTCCACAGCCAAAGCCCGGAAGAGACACAGCGGAGAA 1620
Db 1561 GGGAGCTGCTACGACACCACTCCACAGCCAAAGCCCGGAAGAGACACAGCGGAGAA 1620
QY 1621 GCCCAAGGTCAAGAGATTAATGATGAGCGCACAAGAGAGGGCTGTGTAGAGAGTGC 1680
Db 1621 GCCCAAGGTCAAGAGATTAATGATGAGCGCACAAGAGAGGGCTGTGTAGAGAGTGC 1680
QY 1681 GCAGAGTGGCGGAACATGAGAGACATCTGATCTCTGTGGAGCCTCAATGTTACCT 1740
Db 1681 GCAGAGTGGCGGAACATGAGAGACATCTGATCTCTGTGGAGCCTCAATGTTACCT 1740

Db 1681 GCAGAGTGGCGGAACATTAAGACATCTGATCTCTGTGGAGCCTCAATGTTACCT 1740
QY 1741 GGAACACCCCTCTCTGTTGGAGATGGCCAAATGCAAGAACTGTTCTGAGAG 1800
Db 1741 GGAACACCCCTCTCTGTTGGAGATGGCCAAATGCAAGAACTGTTCTGAGAG 1800
QY 1801 TGCGTACCAAGTACAGAGAGGCTTACAGCTCTACTACGACATCTGCTGGGGGCG 1860
Db 1801 TGCGTACCAAGTACAGAGAGGCTTACAGCTCTACTACGACATCTGCTGGGGGCG 1860
QY 1861 TGAGGTGCTCATGTGCGGAACACACACTGTGACAGTGTCTTGTGGAGTGTGGA 1920
Db 1861 TGAGGTGCTCATGTGCGGAACACACACTGTGACAGTGTCTTGTGGAGTGTGGA 1920
QY 1921 CCTCTTGGTGGGGCGGCTGCCAGGACCATTAAGGAAGACCCCTGGAATGCTA 1980
Db 1921 CCTCTTGGTGGGGCGGCTGCCAGGACCATTAAGGAAGACCCCTGGAATGCTA 1980
QY 1981 CATGTGCGGGCAAGAGGTACCTTACGGCTGTGCGGGCGGAGAGACTGGCCCTCCG 2040
Db 1981 CATGTGCGGGCAAGAGGTACCTTACGGCTGTGCGGGCGGAGAGACTGGCCCTCCG 2040
QY 2041 GCTCCAGATGTTCTTCCGTAATTAACACGACGAGAAATTTGACCTCCAAAGTTTACC 2100
Db 2041 GCTCCAGATGTTCTTCCGTAATTAACACGACGAGAAATTTGACCTCCAAAGTTTACC 2100
QY 2101 ACCTGTCCAGCTGGAAGAGAGCCCATCCGGGTGCTGCTCTTGTGAAGAAATGC 2160
Db 2101 ACCTGTCCAGCTGGAAGAGAGCCCATCCGGGTGCTGCTCTTGTGAAGAAATGC 2160
QY 2161 TACAGGGCTCCTGCTGCTGAAGAGTGTGGGCAATTAAGTGAACCGCTACATTCGTCGA 2220
Db 2161 TACAGGGCTCCTGCTGCTGAAGAGTGTGGGCAATTAAGTGAACCGCTACATTCGTCGA 2220
QY 2221 GGTGTGAGAGACTCATACGCTGGGATGTTGGGACACAGGGGAAGATCATGTAGCT 2280
Db 2221 GGTGTGAGAGACTCATACGCTGGGATGTTGGGACACAGGGGAAGATCATGTAGCT 2280
QY 2281 CGGGAGAGTCCGACAGCTCACAGAAACATTCAGAGAGTGGGGCCATTCGATCTGCT 2340
Db 2281 CGGGAGAGTCCGACAGCTCACAGAAACATTCAGAGAGTGGGGCCATTCGATCTGCT 2340
QY 2341 GATTGGGGGAGTCCCTCAATGACCTCTCATCTCAACCTGCTCCGAAGGGGCTGTA 2400
Db 2341 GATTGGGGGAGTCCCTCAATGACCTCTCATCTCAACCTGCTCCGAAGGGGCTGTA 2400
QY 2401 CGAGGGCACTGGCGGCTCTTCTTGAATTCACGCTCCGTCATGATGCGCGGCCAA 2460
Db 2401 CGAGGGCACTGGCGGCTCTTCTTGAATTCACGCTCCGTCATGATGCGCGGCCAA 2460
QY 2461 GAGGAGATGATGCGCCCTCTTCTGCTCTTGAAGATGTGGGCGCATGGGGGTAG 2520
Db 2461 GAGGAGATGATGCGCCCTCTTCTGCTCTTGAAGATGTGGGCGCATGGGGGTAG 2520
QY 2521 TGACAGAGAGGACATCTGCGATTTCTGAGTGCACCAACCTTGATGATGATGCCAAAGA 2580
Db 2521 TGACAGAGAGGACATCTGCGATTTCTGAGTGCACCAACCTTGATGATGATGCCAAAGA 2580
QY 2581 AGTGTACGTGACACAGAGGCGCGTACTTCTGGGGGTAACTCCCGATGAACAGGCG 2640
Db 2581 AGTGTACGTGACACAGAGGCGCGTACTTCTGGGGGTAACTCCCGATGAACAGGCG 2640
QY 2641 GTTGGCATCCACTGTGAATGATTAACCTGGAGTGAAGAGTGTGAGAGCATGGAGAT 2700
Db 2641 GTTGGCATCCACTGTGAATGATTAACCTGGAGTGAAGAGTGTGAGAGCATGGAGAT 2700
QY 2701 AGCCAAAGTTCAGCAAGTGAAGACCATTAACGAGTCAAGGTCGCAATTAAGCAGGGCAA 2760
Db 2701 AGCCAAAGTTCAGCAAGTGAAGACCATTAACGAGTCAAGGTCGCAATTAAGCAGGGCAA 2760
QY 2761 AGACCAGCATTTCTCTGCTTCATGAATGAGAAAGAGACATCTTATGTGTGACCTGAAT 2820
Db 2761 AGACCAGCATTTCTCTGCTTCATGAATGAGAAAGAGACATCTTATGTGTGACCTGAAT 2820

OY 2821 GGAAGGATTTGGTTCCTCAGTCCACTATCTAGTCCGACATGAGCCGCTTGGC 2880
 Db 2821 GGAAGGATTTGGTTCCTCAGTCCACTATCTAGTCCGACATGAGCCGCTTGGC 2880
 OY 2881 GAGGAGAGAGCTGCTGGCGCGCTGATGAGAGCTGCGGACATCGGCGCTTGGCTC 2940
 Db 2881 GAGGAGAGAGCTGCTGGCGCGCTGATGAGAGCTGCGGACATCGGCGCTTGGCTC 2940
 OY 2941 GCTGAAGAGATTTTGGCTGTGTGTAAGGACATGGGGGCAACTGAGTACGACACA 3000
 Db 2941 GCTGAAGAGATTTTGGCTGTGTGTAAGGACATGGGGGCAACTGAGTACGACACA 3000
 OY 3001 AAGTT 3005
 Db 3001 AAGTT 3005

RESULT 3 AF068625 4192 bp mRNA linear ROD 06-DEC-1999
 LOCUS AF068625 Mus musculus DNA cytosine-5 methyltransferase 3A (Dnmt3a) mRNA,
 DEFINITION complete cds.
 SOURCE AF068625.2 GI:6449467
 KEYWORDS Mus musculus.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 4192)
 AUTHORS Okano, M., Xie, S. and Li, E.
 TITLE Cloning and characterization of a family of novel mammalian DNA
 JOURNAL Nat. Genet. 19 (3), 219-220 (1998)
 MEDLINE 98324766
 PUBMED 9662389
 REFERENCE 2 (bases 1 to 4192)
 AUTHORS Xie, S., Okano, M. and Li, E.
 TITLE Direct Submission
 JOURNAL Submitted (28-MAY-1998) CVRC, Mass. Gen. Hospital, 149 13th Street,
 Charlestown, MA 02129, USA
 REFERENCE 3 (bases 1 to 4192)
 AUTHORS Okano, M., Chijiwa, T., Sasaki, H. and Li, E.
 TITLE Direct Submission
 JOURNAL Submitted (04-NOV-1999) CVRC, Mass. Gen. Hospital, 149 13th Street,
 Charlestown, MA 02129, USA
 REMARK Sequence update by Submitter
 COMMENT On Nov 18, 1999 this sequence version replaced gi:3327977.
 FEATURES
 source
 1. 4192
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /chromosome="12"
 /map="4.0 CM"
 1. 4192
 /gene="Dnmt3a"
 /note="similar to EST sequences deposited in Genbank
 Accession Numbers AA052791, A111043, A1154890, AA240794,
 AA756533, W58898, W59299, W91654, and W91655"
 217. 2943
 /gene="Dnmt3a"
 /function="de novo DNA methylation"
 /note="contains Cys-rich region"
 /codon_start=1
 /product="DNA cytosine-5 methyltransferase 3A"
 /protein_id="AAC40177.2"
 /db_xref="GI:6449468"
 /translation="MPSGPGDTSSSSIEREDDRKEGEQENRNGKEERQEPSATARK
 VGRGRKRHPPESSDTPKPAVTTKSGPMAODSGPDLIPNGDLEKRSBPPEGS
 PMAQKGAPAEGETETPEASRAVENGCCVETKRGASAGEGKQKOTNLSMKME

GSGRRLGGLGWESSLRORPMDLTPQADDPYVTSKREBDMLAPRKREAKKAVYA
 VMNVEPNQSGSOKVEAPASPDAPQPTDAPASVATTPDPPGACGACNATKADQ
 EPEDEDGRCIGEDLVNKLRTFSNMRGRVYSMMNTGRSRAEGTRKRVYFGDGKFSV
 CVETEDLPASFSAHQATINQKQPTKRLATLEVLQVASSAGLTPACHDSDESG
 KAVERYKQNTENALGFRQSGGLPEPEEKNFKEVITDMVETLAPPA
 KPRKSTTERKVEEIDERTRELVTEYRQCRNIEDICSGSLVTLHPLELIG
 KQCKNCFLECATQDDIDYOSTCTCCGAEVLMCGNNMCRCEVCDLVLP
 AAQAIREDPMNCTMCGHGTGTLRRRDMPSRLQMFANNDQEPDPKVPVPA
 EKRKPIVLSLFDGIATGLVLKLDIGLIVDYLASEVCEESTIVGAWRQGLIMVGD
 VRSDRPFHLEENVMAGVSDKRLISRLSEPNVIMDAKEVASAHARFEMNLQPM
 RPLASTVNDKLEIECLHGRIRAKFSKVTITTRSNISIKQKQDHPFPMNEKEDILW
 CTMEERVFQFPVHYTDVSNMSRLARQRLIGRWSVPVIRHLFAPLEKFACV"
 BASE COUNT 1094 a 1072 c 1186 g 840 t
 ORIGIN

Query Match 67.1%; Score 2879.2; DB 10; Length 4192;
 Best Local Similarity 83.8%; Pred. No. 0;
 Matches 3579; Conservative 0; Mismatches 578; Indels 113; Gaps 24;

OY 24 CGGGCGGGCGGACCGCCACCGGCGCATAGCGTACGATCGAACCGCCACCGACGGCT 83
 Db 17 CCGGGCGGGCGGACCGCCACCGGCGCATAGCGTACGATCGAACCGCCACCGACGGCT 76
 OY 84 GACAGAGGACCGCTTCCACGAGGGCTCAACACCGGATCTATGTTTAAGTTAACTCT 143
 Db 77 G-----CACTTTTCGAGGGGCTTGACATCAGGGCTATGTTTAAGTTAACTCT 126
 OY 144 CGCTTCGAAGACGACGANTATCTCTTCCCAAGCCG-AGCAGCCCGCCAGC-CCGCG 201
 Db 127 TCGTTACAAACACGACGCAATTCCTCTGTAAGCCCTCCACCGCCACCGGCGCTCG 186
 OY 202 CAGCCCCAGCTGCTCTCCGCGCGCCAGATGCGCCGCTCCAGCGCGCCCGGGA 261
 Db 187 CAGCCCCAGCTGCTCTCCGCGCGCCAGATGCGCCGCTCCAGCGCGCCCGGGA 239
 OY 262 CACGAGAGCTCTCTCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 321
 Db 240 CACGAGAGCTCTCTCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 299
 OY 322 GCGCGTGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 381
 Db 300 GACCGTGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 359
 OY 382 TGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 441
 Db 360 TGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 419
 OY 442 GATCTCCAAAGTCCCATCCATGAGCCGAGAGCTCAGAGGCGCTCAGAGCTATTAACCAATGG 501
 Db 420 GACCAACCAAGTCTTAGGCCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 479
 OY 502 GCACTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 561
 Db 480 ACACTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 539
 OY 562 GGGCGGGCGGACCGCCACCGGCGCATAGCGTACGATCGAACCGCCACCGACG 621
 Db 540 GGGTGGGCGGACCGCCACCGGCGCATAGCGTACGATCGAACCGCCACCGACG 596
 OY 622 AGTGGAAATGGCTGCTGCGACCGCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 681
 Db 597 TGTGAGAAATGGCTGCTGCTGACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 656
 OY 682 AGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 741
 Db 657 AGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 716
 OY 742 GGTGGCTTGGGCTGGAGTCCAGGCTCCGTAGCGGCGCATGCGGAGGCTCACTTCCA 801
 Db 717 AGGTGGCTTGGGCTGGAGTCCAGGCTCCGTAGCGGCGCATGCGGAGGCTCACTTCCA 776
 OY 802 GCGCGGGGACCCCTACTACATGACGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 861

Db 777 GCGAGGGAGCCCTACTACTATCAAGCAAGCAAGGAGATGAGTGGCTGGACGTTGGAA 836
QY 862 AAGGAGGCTGAGAGAAAGCAAGCTCATTTGACAGCATATATGCTGGAGAGAAACA 921
Db 837 AAGGAGGCTGAGAGAAAGCAAGCTCATTTGACAGCATATATGCTGGAGAGAAACA 896
QY 922 GGGGCCCCGGGAGGCTCAGAAAGGAGGAGGAGCCGCTGCTGCTGGAGAGCCGAC 981
Db 897 GGGCTCTGGAGAGCTCTAGAAAGGTGGAGAGGAGCCAGCTGCTGCTGGAGAGCCGAC 956
QY 982 TGAACCCGACCTCCCACTGCTGCTACACAGCTGAGCCGCTGGGAGCTGGGAGTGGGA 1041
Db 957 GGAACCCGCTTCTCCGAGCTGGAGCAACCCCTGAGCCAGTATGAGAGGAGATGCTGGGA 1016
QY 1042 CAAGATGCGCAACCAAGAGCCATGACAGCCAGCTACAGAGAGCGCGGGCTTTGG 1101
Db 1017 CAAGAAATGCTACCAAAAGCACCCAGCATGAGCTGATGAGATGAGGATGCGGGCTTTGG 1076
QY 1102 CATTTGGGAGCTGCTGTTGGGGAACCTGGGGGCTTCTGCTGGGAGCCGACATTTG 1161
Db 1077 CATTTGGAGAGCTGCTGTTGGGGAACCTTTGGGGCTTCTGCTGGGAGCCGACATTTG 1136
QY 1162 GCTTGGTGGATGACGGGCGGAGCCGAGCAGCTGAAAGGCAACCCGCTGGGCTCATGCTT 1221
Db 1137 GCTTGGTGGATGACAGGCGGAGCCGAGCAGCTGAAAGGCAACCTGCTGGGCTCATGCTT 1196
QY 1222 CGGAGAGCGCAATTTCTCATGCTGCTGTTGAGAGAGCTATGCTGAGCTGCTTTTG 1281
Db 1197 CGGAGAGCGCAATTTCTCATGCTGCTGTTGAGAGAGCTATGCTGAGCTGCTTTTG 1256
QY 1282 CAGTGGCTTCCACAGCGCCAGCTACAAAGAGCCAGTATGAGCGCAAGGCTCATGCA 1341
Db 1257 CAGTGGCTTCCACAGCGCCAGCTACAAAGAGCCAGTATGAGCGCAAGGCTCATGCA 1316
QY 1342 GGTCTGTCAGAGTGGGCGAGCCGCGGGGAAAGCTTTCCGGTGTGCCAGCAGCAGA 1401
Db 1317 AGTCTCTCAGAGTGGGCGAGCCGCGGGGAAAGCTTTCCAGCTTGCATGACATGTA 1376
QY 1402 TGAAGAGTACATTCGCAAGGCGCTGAGAGTGCAGAAACAGCCATGATTTGATGGCCCT 1461
Db 1377 TGAAGAGTACATTCGCAAGGCGCTGAGAGTGCAGAAACAGCTGATTTGATGGCCCT 1436
QY 1462 GGGGCGCTTCCAGGCTTCTGAGCCCTAAGGGGCTGAGCCAGCAGAGAAAGAGAAATTC 1521
Db 1437 CGGAGGCTTCCAGGCTTCTGAGCCCTAAGGGGCTGAGCCAGCAGAGAAAGAGAAATTC 1496
QY 1522 CTACAAAGAAATGACAGGAGCATGTTGGAACCTGAGGAGCAGCTTACGCAACCC 1581
Db 1497 TTACAAAGAAATGACAGGAGCATGTTGGAACCTGAGGAGCAGCTTACGCAACCC 1556
QY 1582 TTCAACCGCAAAAGGCGCGGAGAGACACAGCGGAGAAAGCCAGGCTCAAGAGATAT 1641
Db 1557 CCCACCGCAAAAGGCGCGGAGAGACACAGCGGAGAAAGCCAGGCTCAAGAGATAT 1616
QY 1642 TGAATGACGCAAGAGAGGCGCTGATGACAGGCTGCGGAGAAAGTGCAGAACTTGA 1701
Db 1617 TGAATGACGCAAGAGAGGCGCTGATGATGAGGCTGCGGAGAAAGTGCAGAACTTGA 1676
QY 1702 GGACATCTGATCTCTGTTGGAGCCTCAATGTTACCTGGAAACCCCTCTTCTGTTG 1761
Db 1677 GGACATCTGATCTCTGTTGGAGCCTCAATGTTACCTGGAAACCCCTCTTCTGTTG 1736
QY 1762 AGGAATGTGCCAAACCTGCAAGAACTGTTCTGAGATGTGCTACCAAGTACGACAGCA 1821
Db 1737 AGGAATGTGCCAAACCTGCAAGAACTGTTCTGAGATGTGCTACCAAGTACGACAGCA 1796
QY 1822 CGGCTACAGCTCCTACTGACACCATCTGCTGGGGGCGTGAAGTGTCTATGTCGGAAA 1881
Db 1797 TGGGTACAGCTCCTACTGACACCATCTGCTGGGGGCGTGAAGTGTCTATGTCGGAAA 1856
QY 1882 CAACAACTGCTGACAGTGTCTTTGCTGAGAGTGTGAGACCTTGGTGGGGCGGGGCG 1941
Db 1857 CAACAACTGCTGACAGTGTCTTTGCTGAGAGTGTGAGACCTTGGTGGGGCGGGGCG 1916
QY 1942 TGCCCGAGCAGCATTAAGAAAGACCCCTGGAAGCTGTACATATGTCGGGCGCAAGGCTAC 2001
Db 1917 TGCTCAGCAGCAGCATTAAGAAAGACCCCTGGAAGCTGTACATATGTCGGGCGTAAGGGCAC 1976
QY 2002 CTAGGGGCTCTGGGGGGGAGAGAGCTGAGCCCTCCGGCTCCAGATGTTCTTCTGCTAA 2061
Db 1977 CTATGAGCTCTGCAAGAGAGGAGAGCTGAGCCCTCTCTGAGCTCAGATGTTCTTGGCAA 2036
QY 2062 TAAACAGCAGCAGAAATTTAGACCTCCCAAGGTTTACCACTGTCCACCTGAGAAAGAG 2121
Db 2037 TAAACAGCAGCAGAAATTTAGACCTCCCAAGGTTTACCACTGTCCACCTGAGAAAGAG 2096
QY 2122 GAAGCCATCCGGGTGCTGCTCTTGTGATGATGATGATGATGATGATGATGATGATGAT 2181
Db 2097 GAAGCCATCCGGGTGCTGCTCTTGTGATGATGATGATGATGATGATGATGATGATGAT 2156
QY 2182 GGAATTTGGGATTCAGGCTGAGCAGCTATGCTGAGGAGTGTGAGAGCTTCAATCAC 2241
Db 2157 GGAATTTGGGATTCAGGCTGAGCAGCTATGCTGAGGAGTGTGAGAGCTTCAATCAC 2216
QY 2242 GGTGGGATGCTGGGCGCAGCAGGGAAGATCATGTACGTGCGGAGCTGCCAGGCTCAC 2301
Db 2217 GGTGGGATGCTGGGCGCAGCAGGGAAGATCATGTACGTGCGGAGCTGCCAGGCTCAC 2276
QY 2302 ACAGAACATATCCAGAGATGAGGAGCCATTCGATCTGCTGATATGAGGAGCTCCCTCAA 2361
Db 2277 ACAGAACATATCCAGAGATGAGGAGCCATTCGATCTGCTGATATGAGGAGCTCCCTCAA 2336
QY 2362 TGAACCTTCCATGCTCAACCTGCTGCAAGGAGCTTACAGAGGCTGCTGAGGAGCTTCA 2421
Db 2337 TGAACCTTCCATGCTCAACCTGCTGCAAGGAGCTTACAGAGGCTGCTGAGGAGCTTCA 2396
QY 2422 CTTTGAATTTACCGCTCTCTGATGATGATGATGATGATGATGATGATGATGATGAT 2481
Db 2397 CTTTGAATTTACCGCTCTCTGATGATGATGATGATGATGATGATGATGATGATGAT 2456
QY 2482 CTTTGGCTCTTTGAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 2541
Db 2457 CTTTGGCTCTTTGAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 2516
QY 2542 ATTTCGAGTCCAAACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2601
Db 2517 ATTTCGAGTCCAAACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2576
QY 2602 CCGCTACTTCTGGGTAACCTTCCGATGATGATGATGATGATGATGATGATGATGATGAT 2661
Db 2577 CCGCTACTTCTGGGTAACCTTCCGATGATGATGATGATGATGATGATGATGATGATGAT 2636
QY 2662 TAACTGAGGCTGAGAGAGTGTGAGAGCATGAGAGATGAGAGATGAGAGATGAGAG 2721
Db 2637 TAACTGAGGCTGAGAGAGTGTGAGAGCATGAGAGATGAGAGATGAGAGATGAGAG 2696
QY 2722 GACCATTAACGAGTCAAACTTCAATTAAGCAGGAGGCAAGCAGCATTTCTGCTT 2781
Db 2697 GACCATTAACGAGTCAAACTTCAATTAAGCAGGAGGCAAGCAGCATTTCTGCTT 2756
QY 2782 CATGATGAGAAAGAGACATCTTATGAGTGCATGAAATGAGAAAGGATATTTGCTTCC 2841
Db 2757 CATGATGAGAAAGAGACATCTTATGAGTGCATGAAATGAGAAAGGATATTTGCTTCC 2816
QY 2842 AGTCCATTAATACGCTTCCCAATGAGAGCCCTTGGCAGAGAGAGAGAGAGAGAGAG 2901
Db 2817 AGTCCATTAATACGCTTCCCAATGAGAGCCCTTGGCAGAGAGAGAGAGAGAGAGAG 2876
QY 2902 GTCATGAGGCTGTCAGTACCTCCGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2961
Db 2877 ATGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2936
QY 2962 TGTGTAAGGAGCATGAGGAGCAAACTGAGAGTACGCA---CACAAAGTTAAACAAAC 3017
Db 2937 TGTGTAAGGAGCATGAGGAGCAAACTGAGAGTACGCA---CACAAAGTTAAACAAAC 2996

VGRPKRKHHPVESSDTPKDPVATTKSQPMADSGPDLLENGDLEKRSPOPEBS
 PAQGGKGPAGEGETPPEFASRAVENCCYTKBGRGASABEGKOKTNETSKME
 GSRGRGLGESSLROBPMRLTPOADDPYTSKRDEMLARKRAEKKATYIA
 VYNAVBNASSESOKEVASPPAQQPTDPAISPYVATTPPEVGGDAGKNAFKAD
 BEYEDRGFEIGELVWCKLRFSWMPKIVSMWMTGRSRAAEGTWMVMDKRSY
 VYVEKLMPLSSFCSAFHQATYKNOBKRAIYLVQVSSRAGKLEPACHDEDSG
 KAVEYONKMIEMALGGFQPSGPKGLEPEEKBNPKYEYTMWVPEFAAAPP
 KPEKSTTEKPKVKEIIDERTREBLVEYRKRNRNEDICISGSLNVLLEHPLFG
 MCONKCEICAYODDGYOCTYCCGGRBVLMMGNKRCRCVCYDLVPSG
 AAQAKIKEDPNKCYMGHKGTGLLRREDRPSRLQMFANNDHDFDPKYPVPA
 EKRPTRVLSLDGLATGLVLDLGIQDRIASEVCEDSITVGVHRQKIMVYD
 VRSYVQKHIQEWPPDLYIGSPCNDLSIVNPARGLVDAKESVAFHRLHDAPK
 EGDDEPFELWENYVAMGVSDKDISRLESNPMDAKESVAHRAEPKMLAPDM
 RLASTVNDKLELOCELEHRIAKESKVTITTRSNISIOGDQHPVPMNEEDILM
 CTMEHVEFEPHYTDVSNMSRLAROLLGRSMVSVIHLPLAKEYRACY*

BASE COUNT 1067 a 1057 c 1164 g 806 t

Query Match 65.8%; Score 2825.4; DB 10; Length 4094;

Best Local Similarity 84.1%; Pred. No. 0;

Matches 3487; Conservative 0; Mismatches 546; Indels 114; Gaps 22;

1 GCCCGGCGACAGAGGCGCGAGCCGCCGCCAGCCACCAGCCATPAGGTGAGCC 60
 26 GCCCGGCGACAGAGGCGAG---CCGGCGCGCCGACCCGCCGCCACAGCGCAAGCC 81
 61 ATCGAAGCCGCCACGAGGCTGACAGAGGCGACCGTTACACAGAGGCTCAACACCGG 120
 82 GCGTGAAGCCAGCGGTGAGGCTG-----CACTTTCCGAGGGCTGACATCAGG 131
 121 ACTATATTTAAATTTAACTCTCGCTCCAAAGACAGATTAATCTTCCCAAGCC 180
 132 GTCTATGTTTAACTTACTCTCTGTTACAAAGACAGCGCAATTCCTTCTGAGCC 191
 181 C-AGCAGCCCGCAGC-CCCGCGCAGCCCGCCTCTCCCGCGCGCAGATGCCGCC 238
 192 CTCGCGCCCGCAGCGCCCTCCAGCCCGCTCCCTACTGCCA-----GCA 244
 239 ATGCGCTCCAGCGCGCGCGCGGACACAGAGCTGCTGCGGAGGAGGAGAGCGCA 258
 245 ATGCGCTCCAGCGCGCGCGCGGACACAGAGCTCTCTGAGGCGGAGGTGATCGA 304
 299 AAGCAGGAGAGAGAGAGAGCGCGCTGCGCAAGAGAGAGCGCCAGCCAGCCACAC 358
 305 AAGGAGAGAGAGAGAGAGAGAGAGCGGTGCGCAAGAGAGAGCGCCAGCGCCAGCC 364
 359 ACGGACAG 418
 365 ACGGCGCGGAG 424
 419 GACAGCGCAAG 478
 425 GACAGCGCAAG 484
 479 GCGCTCAGAGCTATTTACCAATGGGAGCTTGGAGAGAGAGAGAGAGAGAGAGAG 538
 485 CCTCAGAGCTGTGCTACCAATGGGAGCTTGGAGAGAGAGAGAGAGAGAGAGAG 544
 539 GGGAGCGCTGTGGGAG 598
 545 GGGAGCGCTGTGGGAG 601
 599 ACCCTGCGTGAAG 658
 602 ACCCTGCGTGAAG 661
 659 GAGAGCGCTGTGGGAG 718
 662 GAGAGCGCTGTGGGAG 721
 719 GAGAGCGCTGTGGGAG 778
 722 GAGAGCGCTGTGGGAG 781

779 CCCATGCGAGGCTGACCTTCCAGGCGCGGGGAGCCCTTACTACATCAGCAAGCGAG 838
 782 CCCATGCGAGGCTGACCTTCCAGGCGCGGGGAGCCCTTACTACATCAGCAAGCGAG 841
 839 GAGAGTGCCTGCGACAGCTGCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 898
 842 GATGATGCTGCGACAGTGTGAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 901
 899 ATGATGCTGCTGAG 958
 902 ATGAGAGTGTGAG 961
 959 CCGCTGCTGTGAG 1018
 962 CCGCTGCTGTGAG 1021
 1019 CCGCTGCGGAG 1078
 1022 CCGAGTGAAG 1081
 1079 TACGAG 1138
 1082 TATGAG 1141
 1139 TCGTGTGAG 1198
 1142 TCGTGTGAG 1201
 1199 GGCAGCCGCTGAG 1258
 1202 GGCAGCTGCTGAG 1261
 1259 CTGATGCGCTGAG 1318
 1262 CTGATGCGCTGAG 1321
 1319 ATGATGCGAG 1378
 1322 ATGATGCGAG 1381
 1379 TTCCGCTGTGAG 1438
 1382 TTCCGCTGTGAG 1441
 1439 AAGCCATGATGATGAG 1498
 1442 AAGCCATGATGATGAG 1501
 1499 CCACGAG 1558
 1502 CCACGAG 1561
 1559 GAGCAGAGCTGCTACAG 1618
 1562 GAGCAGAGCTGCTACAG 1621
 1619 AAGCCATGATGATGAG 1678
 1622 AAGCCATGATGATGAG 1681
 1679 CCGCAG 1738
 1682 CCGCAG 1741
 1739 CTGAG 1798
 1742 CTGAG 1801
 1799 TGTGCGTACAGTACAG 1858
 1802 TGTGCGTACAGTACAG 1861

Db	1201	CCCGAGTCCATTACTAGACGTGCTCCAACTGAGACGGCTTGGGAGAGGACGAGACTGCTGGG	1269
OY	2899	CGCGTCATGGAGGCTGCCGATCATCCGACACCTCTGCTGCTCGCGTGAAGAGATATTTTCG	2956
Db	1261	CCGGTCATGGAGGCTGCCGATCATCCGACACCTCTGCTGCTCGCGTGAAGAGATATTTTCG	1320
OY	2959	GTGTGTGTAAGGACATGGGGGCAACTGAGCTAGCGACACAAAGTTAAACAAACAGACA	3018
Db	1321	GTGTGTGTAAGGACATGGGGGCAACTGAGCTAGCGACACAAAGTTAAACAAACAGACA	1380
OY	3019	AAAAACCAAAACATATATTAACACACCAAGACATGAGATGAGAGAAATATACGACCC	3078
Db	1381	AAAAACCAAAACATATATTAACACACCAAGACATGAGATGAGAGAAATATACGACCC	1440
OY	3079	AGAGAGAAAAAGAAATTTTAAACAAAAACACAGAGGGGAAATACCGGAGGGCTTTTCG	3138
Db	1441	AGAGAGAAAAAGAAATTTTAAACAAAAACACAGAGGGGAAATACCGGAGGGCTTTTCG	1500
OY	3139	CTTGCAGAAAAGGTTTGACATCATCTCCTGATTTTTCATGTATTTCTGACGCTATTT	3198
Db	1501	CTTGCAGAAAAGGTTTGACATCATCTCCTGATTTTTCATGTATTTCTGACGCTATTT	1560
OY	3199	AAAAACAAAAACCAAGCTCCCTCCCTTCCTCCGCCCTCCCTTTTTCGCTAGACCTT	3258
Db	1561	AAAAACAAAAACCAAGCTCCCTCCCTTCCTCCGCCCTCCCTTTTTCGCTAGACCTT	1620
OY	3259	TTATTTTCTACTGTTTTCAGAGGGGTTTTCTGTTTTGGGTTTTGTTTCTTGCTGCTGA	3318
Db	1621	TTATTTTCTACTGTTTTCAGAGGGGTTTTCTGTTTTGGGTTTTGTTTCTTGCTGCTGA	1680
OY	3319	CTGAAACAGAAAGCTATTTGCGACGAAAAATCACTAACAAAAATATATATATACCTTCG	3378
Db	1681	CTGAAACAGAAAGCTATTTGCGACGAAAAATCACTAACAAAAATATATATATACCTTCG	1740
OY	3379	AGAGAAAGAGGTGGGAGAGAGAGAAAAAGGAAATTTTAAAGAAATCTATATATGGGT	3438
Db	1741	AGAGAAAGAGGTGGGAGAGAGAGAAAAAGGAAATTTTAAAGAAATCTATATATGGGT	1796
OY	3439	TGTTTACTATATCTT	3498
Db	1797	TGTTTACTATATCTT	1856
OY	3499	TTTTTTTGTCTCTTAGCTGATCGATAGATAGACACACAGAGGGAGCGAAAGAGAGA	3558
Db	1857	TTTTTTTGTCTCTTAGCTGATAGATAGAGAACACACAGAGGGAGCGAAAGAGAGA	1916
OY	3559	CACTCAGGCGCGACGATTCCTCCACACCACTAGCTGTGCTGCACACACATTTCTGTG	3618
Db	1917	CACTCAGGCGCGACGATTCCTCCACACCACTAGCTGTGCTGCACACACATTTCTGTG	1976
OY	3619	CACCGAAAAAGAACCCAGATTAGCAGAGAGGAGACAGAGAACACACCAAGACATTTTTC	3678
Db	1977	CACCGAAAAAGAACCCAGATTAGCAGAGGAGACAGAGAACACACCAAGACATTTTTC	2036
OY	3679	TACAGTATTTTCAGATGCCCTACACACAGAGAAACCTTGAAGAAATCAGTTTCTAGAGGC	3738
Db	2037	TACAGTATTTTCAGATGCCCTACACACAGAGAAACCTTGAAGAAATCAGTTTCTAGAGGC	2096
OY	3739	GCTGTACACCTCTGTTTACAGTTTATATATATATATATATATATATATATATATATATAA	3798
Db	2097	GCTGTACACCTCTGTTTACAGTTTATATATATATATATATATATATATATATATATATAA	2156
OY	3799	GGTACTGTTTAACTACTGTATATAA	2178
Db	2157	GGTACTGTTTAACTACTGTATATAA	2178
RESULT 6			
LOCUS	BC018214	1758 bp	mRNA
DEFINITION	Homo sapiens, clone IMAGE:3862699,	mrna,	partial cds.
ACCESSION	BC018214		
VERSION	BC018214.1	GI:17390481	

KEYWORDS	Homo sapiens.
SOURCE	Homo sapiens
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE	1 (bases 1 to 1758)
AUTHORS	Direct Submission
JOURNAL	Submitted (03-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: http://mgc.ncl.nih.gov
COMMENT	Contact: MGC help desk Email: cgsaps.rem@ncl.nih.gov Tissue Procurement: DCTD/DTP CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: http://www.hgsc.bcm.tmc.edu/cdna/ Contact: villalona@bcm.tmc.edu. Villation, D.K., Luna, R.A., Hale, S.M., Huijck, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B., Hodgson, A., Bouek, J., Yu, W., Muzny, D.M., Gibbs, R.A.
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov Series: IRAP plate: 21 Row: b Column: 23. Location/Qualifiers 1..1758 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:3862699" /tissue_type="Ovary; adenocarcinoma" /clone_lib="NHL/MGC_66" /lab_host="DH10B" /note="Vector: pCMV-Sport6" <1..859 /codon_start=2 /product="Unknown (protein for IMAGE:3862699)." /protein_id="AAH8214.1" /db_xref="GI:17390482" /translacion="AEKRKPIRVLSLPDGIALGLVLKLDIGVDYRTASEYCEDSITVGVARHCKIMYGVDSVTKHIDQEPFDLVIGGSPCHNDLSIVNPARKIYLGETGRALHFARFPMGNRLPGRMRDLASTVNDKLELOECLNGRILAKSKVRITTRSSINIOGGDQHPPVARNKKEDLIWCTEMERVEGFVIHYTVDSNMSLARORLLGSWSVIPRIHLFAPLKEKYACY"
CDS	
BASE COUNT	513 a 376 c 428 g 441 t
ORIGIN	
Query Match	38.8%; Score 1667; DB 9; Length 1758;
Best Local Similarity	99.2%; Pred. No. 0;
Matches 1697; Conservative	0; Mismatches 10; Indels 4; Gaps 2;
Oy	2110 AGCGAGAAGAGAGCCATCGCGGTCGTCTCTTGATGGATGCCACAGGGCT 2169
Db	1 AGCTGATAAAGAGAACCCATTCGGGTCGTCTCTTTATGGAATGCCACAGGGCT 60
Oy	2170 CCTGTGCTTAAGACTTGGCATTCAGGTGGACCGCTCAATTGCCTCGAGAGTGTGCA 2229
Db	61 CCTGTGCTTAAGACTTGGCATTCAGGTGGACCGCTCAATTGCCTCGAGAGTGTGCA 120
Oy	2230 GAACCTCATTACGGTGGCATGTGTGGCCGACCACAGGGGAAATCATGTACGTGGGACGT 2289
Db	121 GAACCTCATTACGGTGGCATGTGTGGCCGACCACAGGGGAAATCATGTACGTGGGACGT 180
Oy	2290 CCGAGAGCTCACACAGACATATCCAGAGTAGTGGGCCCATTCGATCTGGTGATTTGGGGG 2349
Db	181 CCGAGAGCTCACACAGACATATCCAGAGTAGTGGGCCCATTCGATCTGGTGATTTGGGGG 240

[illegible][illegible]

Center project name: H_NH0458N05

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P. Y., Zhou, B., Firengen, E., Teleno, M., Caranese, J. J. and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RPI1-509E16; the clone sequenced to the right is RPI1-579I5. Actual start of this clone is at base position 1 of RPI1-458N5; actual end is at base position 197652 of RPI1-579I5.

Ambiguous base exists at 127455. Single plasmid region exists between 126594 and 126656. Data from AC104659 was used to finish AC012074.

FEATURES

source

Location/Qualifiers
1. 197652
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RPI1-458N5"
/clone.lib="RPCI-11"
98. 161
/rpt_family="GA-rich"
repeat_region
1121. 1272
/rpt_family="(TA)n"
repeat_region
1326. 1625
/rpt_family="Alu"
repeat_region
1664. 1699
/rpt_family="AT-rich"
repeat_region
1992. 2266
/rpt_family="Alu"
repeat_region
2287. 2443
/rpt_family="MIR"
repeat_region
4014. 4143
/rpt_family="Alu"
repeat_region
4144. 4454
/rpt_family="Alu"
repeat_region
4455. 4657
/rpt_family="Alu"
repeat_region
4695. 4781
/rpt_family="L1"
repeat_region
4782. 5086
/rpt_family="Alu"
repeat_region
5099. 5326
/rpt_family="Alu"
repeat_region

repeat_region 5336. 5589
/rpt_family="L1"
repeat_region 5815. 5955
/rpt_family="MIR"
repeat_region 6009. 6320
/rpt_family="Alu"
repeat_region 6574. 6726
/rpt_family="L2"
repeat_region 6951. 7139
/rpt_family="MIR"
repeat_region 7378. 7481
/rpt_family="MIR"
repeat_region 7508. 7689
/rpt_family="L2"
repeat_region 7701. 7946
/rpt_family="Alu"
repeat_region 7947. 8214
/rpt_family="L2"
repeat_region 8215. 8515
/rpt_family="Alu"
repeat_region 8516. 8538
/rpt_family="L2"
repeat_region 8539. 8667
/rpt_family="L2_type"
repeat_region 8668. 8868
/rpt_family="L2"
repeat_region 8869. 9176
/rpt_family="Alu"
repeat_region 9177. 9406
/rpt_family="L2"
repeat_region 9460. 9534
/rpt_family="L2"
repeat_region 9660. 9965
/rpt_family="Alu"
repeat_region 9998. 10034
/rpt_family="(TG)n"
repeat_region 10651. 10815
/rpt_family="Alu"
repeat_region 10816. 10835
/rpt_family="(TTG)n"
repeat_region 10836. 11121
/rpt_family="Alu"
repeat_region 11389. 11683
/rpt_family="Alu"
repeat_region 12188. 12529
/rpt_family="Alu"
repeat_region 12975. 13131
/rpt_family="L1"
repeat_region 13167. 13466
/rpt_family="Alu"
repeat_region 13470. 13533
/rpt_family="L1"
repeat_region 13765. 14072
/rpt_family="Alu"
repeat_region 14216. 14235
/rpt_family="(TTTA)n"
repeat_region 14241. 14519
/rpt_family="Alu"
repeat_region 14890. 15178
/rpt_family="CR1"
repeat_region 15717. 15841
/rpt_family="L1"
repeat_region 16187. 16217
/rpt_family="AT-rich"
repeat_region 21273. 21317
/rpt_family="(A)n"
repeat_region 22691. 22992
/rpt_family="Alu"
repeat_region 23576. 23872
/rpt_family="Alu"
repeat_region 24866. 24966
/rpt_family="L1"
repeat_region 24994. 25287

```
repeat_region      26925..27112      /pct_family="Alu"  
repeat_region      28773..28875      /pct_family="MER1_type"  
repeat_region      30336..30532      /pct_family="Alu"  
repeat_region      30533..30823      /pct_family="L1"  
repeat_region      30824..30998      /pct_family="Alu"  
repeat_region      30824..30998      /pct_family="L1"
```

Query Match	32.8%	Score 1408.6;	DB 9;	Length 197652;
Best local Similarity	99.1%;	Pred. No. 5.1e-278;		
Matches 1438;	Conservative 0;	Mismatches 9;	Indels 4;	Gaps 2;

2625	AGGGATTTGGTTTCCAGTCCACACATATACAGAGCTGCCAACATGAGCGCTTTGGGAGG	2884
21930	AGGGTATTTGGTTTCCAGTCCACACTATACAGAGCTGCCAACATGAGCGCTTTGGGAGG	21871
2885	CAGGAGTGTGGGCGGCTCATGAGAGCTGGCCAGTATCCGGCACTCTTTCGCTCCGCG	2944
21870	CAGGAGTGTGGGCGGCTCATGAGAGCTGCCAGTATCCGGCACTCTTTCGCTCCGCG	21811
2945	AAGAGATATTTTTCGCTGTGTAAAGGACATGGGGCCAACTGAGTACGACACAAAGT	3004
21810	AAGAGATATTTTTCGCTGTGTAAAGGACATGGGGCCAACTGAGTACGACACAAAGT	21751
3005	TAAACAACAACAAAAACACAAAACATATATAAACACCAGAACATGAGATGGAGAG	3064
21750	TAAACAACAACAAAAACACAAAACATATATAAACACCAGAACATGAGATGGAGAG	21691
3065	AAGATACAGACCCAGAGAGAAAAAGAAATTTAAAAACAAAACACAGAGGGGGAATA	3124
21690	AAGATACAGACCCAGAGAGAAAAAGAAATTTAAAAACAAAACACAGAGGGGGAATA	21631
3125	CCGAGAGGCTTTCGCTTTCGAAAAAGGTTGGACATCATCTCGATTTTTCATGTTAT	3184
21630	CCGAGAGGCTTTCGCTTTCGAAAAAGGTTGGACATCATCTCGATTTTTCATGTTAT	21571
3185	CTTTCAGTCCATTTTAAAAAACCAAGTCCCTTCCTCTCCCTCCCTCCCTTTT	3244
21570	CTTTCAGTCCATTTTAAAAAACCAAGTCCCTTCCTCTCCCTCCCTCCCTTTT	21511
3245	TTTCGTAGAGACCTTTTATTTTCTACTCTTTTCAGAGGGGTTTCTGTTTGGTTGGGTTT	3304
21510	TTTCGTAGAGACCTTTTATTTTCTACTCTTTTCAGAGGGGTTTCTGTTTGGTTGGGTTT	21451
3305	GTTTCTTGTGTGACTGAACAAGAGGTTATTGACGACAAAATCAGTAAACAAAATAG	3364
21450	GTTTCTTGTGTGACTGAACAAGAGGTTATTGACGACAAAATCAGTAAACAAAATAG	21391
3365	TAAACAATACCTTCAGAGGAAAGGTGGAGAGAGAAAAAGGAAATTTTAAAGAA	3424
21390	TAAACAATACCTTCAGAGGAAAGGTGGAGAGAGAAAAAGGAAATTTTAAAGAA	21334
3425	TCTATATATTTGGGTTGTTTTTTTTTTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	3484
21333	TCTATATATTTGGGTTGTTTTTTTTTTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	21271
3485	TTACTATATATCTTTTTTTTGTGTGCTCTAGCCTGATCAGATAGAGACCAAGCAGGGA	3544
21274	TTACTATATATCTTTTTTTTGTGTGCTCTAGCCTGATCAGATAGAGACCAAGCAGGGA	21211
3545	CGGAAAAGAGAGACATCAGGGGGGACGATTCCTCCGACGACGACGAGCTGTGCGCA	3604
21214	CGGAAAAGAGAGACATCAGGGGGGACGATTCCTCCGACGACGAGCTGTGCGCA	21151
3605	GCACCATTTCCGTGACAGCAAAACAGAACCACTTGAACAGACAGGAGACAGAAACACAC	3664
21154	GCACCATTTCCGTGACAGCAAAACAGAACCACTTGAACAGACAGGAGACAGAAACACAC	21091
3665	ACAAGACATTTTTCATCAGTATTTTTCAGGTGCTTACCAACACAGAGAACTTGAAGAAATC	3724

Dd	21094	ACAAAGACATTTTCTACAGTAITTCAGGTCCTAACCAACAGAGAAAACCTTGGAAGAAATC	2103531
OY	3725	AGTTTCTAGAAGCGCGTGTACCCTCTTGTATTACAGTTATATATATATGATAGATATGAG	3784
Dd	21034	AGTTCTAGAAGCGCGGTACCTGTACCTGTATTACAGTTATATATATATGATAGATATGAG	2097575
OY	3785	ATATATATATAAAGGACTGTATACACTGTACAACCCGACATCTCATATANGGCGTTCA	3844
Dd	20974	ATATATATATAAAGGACTGTATACACTGTACAACCCGACATCTCATATANGGCGTTCA	2091515
OY	3845	AACAGCAGATGAGTAAAAACATCACGTTCCAGTTGCCCTTGCSCAAAGGGTTTCACC	3904
Dd	20914	AACAGCAGATGAGTAAAAACATCACGTTCCAGTTGCCCTTGCSCAAAGGGTTTCACC	2085555
OY	3905	AAGGATGAGAAAGGGAGACAGTTCAGATGCGCGGCTTCACGGTGGGCGTCTCCCT	3964
Dd	20854	AAGGATGAGAAAGGGAGACAGTTCAGATGCGCGGCTTCACGGTGGGCGTCTCCCT	2079355
OY	3965	TGTTTTGTAAACGAAGTAAAGAGAGAGAACTTGGGAGCCACAGTTCCTCCGCCAAAAAGG	4024
Dd	20794	TGTTTTGTAAACGAAGTAAAGAGAGAGAACTTGGGAGCCACAGTTCCTCCGCCAAAAAGG	2073555
OY	4025	GCGTAGATGAGAGTCGTGGGCGCGGTGACACGTCGAGATNGGATTTATCCAGATCATAC	4084
Dd	20734	GCGTAGATGAGAGTCGTGGGCGCGGTGACACGTCGAGATNGGATTTATCCAGATCATAC	2067575
OY	4085	AATAACCCCTTGTATGTTTCTTAAAGGAACACTCCCTCGGCAAGATGGCAGAGAGGTACGG	4144
Dd	20674	AATAACCCCTTGTATGTTTCTTAAAGGAACACTCCCTCGGCAAGATGGCAGAGAGGTACGG	2061515
OY	4145	AGCTTCAGGCCCCAGTTCTCACCTTACGCCAATTCGAGGGCGTCTGTGGTGGGATCACA	4204
Dd	20614	AGCTTCAGGCCCCAGTTCTCACCTTACGCCAATTCGAGGGCGTCTGTGGTGGGATCACA	2055555
OY	4205	ACTAATCCAGAGTGTGGGAAAGTGACAGTCMAAACCCCAACCTTGAGCAATAAAAAAAAA	4264
Dd	20554	ACTAATCCAGAGTGTGGGAAAGTGACAGTCMAAACCCCAACCTTGAGCAATAAAAAAAAA	2049555
OY	4265	TACAAACGTA 4275	
Dd	20494	TACAAACGTA 20484	

AC120824	RESULT 8
LOCUS	AC120824
DEFINITION	AC120824
ACCESSION	AC120824
VERSION	AC120824.2
KEYWORDS	HTG: HTGS-PHASE1.
SOURCE	Rattus norvegicus
ORGANISM	Rattus norvegicus
AC120824	123936 bp
Rattus norvegicus clone CH230-282016, ***	DNA linear HTG 23-JUL-2002
***, 38 unordered pieces.	SEQUENCING IN PROGRESS
AC120824.2	GI:21902541
HTG: HTGS-PHASE1.	
Rattus norvegicus	
Rattus norvegicus	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;	
Rattus.	
Musny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,	
Alsbrooks, S.L., Amaralunga, H.C., Are, J.R., Ayale, M., Banks, T.,	
Barbala, J., Benton, J., Bimege, K., Blankenburg, K., Bonnin, D.,	
Bouck, J., Bowie, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P.,	
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,	
Carroll, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,	
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,	
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,	
David, M.L., Davis, C., Day-Carroll, L., Dederich, D.A.,	
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinu, H.H.,	
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,	
Earhart, C., Edgar, D., Edwards, C.C., Elhai, C., Escotto, M.,	
Falls, T., Ferrnuto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,	
Gabis, A., Gao, J., Garcia, A., Garner, T., Gara, N., Gill, R.,	
Gorell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,	

Harris, C., Harris, K., Hart, M., Haylak, P., Hawes, A., Hernandez, J.,
 Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B.,
 Homel, P., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
 Jacobson, B., Jia, Y., Johnson, R., Joliver, S., Joudan, S.,
 Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
 Kravovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
 Li, J., Li, Z., Licharge, O., Liu, C., Liu, J., Liu, W., Louisedge, H.,
 Lozada, R.U., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
 Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
 Massey, E., Mathney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M.,
 Miner, G., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
 Nguyen, N., Nickerson, E., Nwokwenkwo, S., Ogun, M., Okunolu, C.,
 Ogunyemi, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
 Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
 Rives, M., Scott, G., Shen, H., Shooshari, N., Sisson, I.,
 Scheraga, S., Sonalke, T., Sparks, A., Stanley, H., Stone, H.,
 Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
 Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
 Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
 Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,
 Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstein, G. and Gdbbs, R.
 Direct Submission
 Unpublished
 2 (bases 1 to 123936)
 Morley, K.C.
 Direct Submission
 Submitted (13-MAY-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 123936)
 Worley, K.C.
 Direct Submission
 Submitted (23-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jul 18, 2002 this sequence version replaced gi:20531783.
 ----- Genome Center -----
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 Project Information
 Project name: CH230-282D16
 Center clone name: CH230-282D16
 ----- Summary Statistics -----
 Sequencing vector: Plasmid;
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 95907 bases at least Q40
 Consensus quality: 99223 bases at least Q30
 Consensus quality: 101036 bases at least Q20
 ----- NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_dirft_data.html).
 NOTE: This is a 'working draft' sequence. It currently
 consists of 38 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.
 * 1010 1009: contig of 1009 bp in length
 * 1110 2113: gap of unknown length
 * 2114 2213: gap of unknown length
 * 2214 3233: contig of 1020 bp in length
 * 3234 3333: gap of unknown length
 * 3334 4442: contig of 1109 bp in length
 * 4443 4542: gap of unknown length
 * 4543 6202: contig of 1660 bp in length
 * 6203 6302: gap of unknown length
 * 6303 7464: contig of 1162 bp in length
 * 7465 7565: gap of unknown length
 * 7566 8772: contig of 1208 bp in length
 * 8773 8872: gap of unknown length
 * 8873 10371: contig of 1499 bp in length
 * 10372 10471: gap of unknown length
 * 10472 11714: contig of 1243 bp in length
 * 11715 11814: gap of unknown length
 * 11815 12924: contig of 1110 bp in length
 * 12925 13024: gap of unknown length
 * 13026 14347: contig of 1323 bp in length
 * 14348 15653: gap of unknown length
 * 15654 15753: gap of unknown length
 * 15754 17796: contig of 2044 bp in length
 * 17797 17896: gap of unknown length
 * 17897 18937: contig of 1041 bp in length
 * 18938 19037: gap of unknown length
 * 19038 21543: gap of 2506 bp in length
 * 21544 21644: gap of unknown length
 * 21645 24541: contig of 2857 bp in length
 * 24542 24640: gap of unknown length
 * 24641 28112: contig of 3472 bp in length
 * 28113 28212: gap of unknown length
 * 28213 31101: contig of 2889 bp in length
 * 31102 31202: gap of unknown length
 * 31203 34415: contig of 3214 bp in length
 * 34416 34515: gap of unknown length
 * 34516 37353: contig of 2838 bp in length
 * 37354 37454: gap of unknown length
 * 37455 40939: contig of 3485 bp in length
 * 40940 41039: gap of unknown length
 * 41040 44331: contig of 3293 bp in length
 * 44332 44431: gap of unknown length
 * 44432 48035: contig of 3604 bp in length
 * 48036 48135: gap of unknown length
 * 48136 51181: contig of 3046 bp in length
 * 51182 51281: gap of unknown length
 * 51282 55361: contig of 4080 bp in length
 * 55362 55461: gap of unknown length
 * 55462 59580: contig of 4119 bp in length
 * 59581 63525: gap of unknown length
 * 63526 63625: contig of 3845 bp in length
 * 63626 67434: gap of unknown length
 * 67435 72194: contig of 3809 bp in length
 * 72195 72194: gap of unknown length
 * 72195 72294: contig of 4660 bp in length
 * 72295 77777: gap of unknown length
 * 77778 81340: contig of 5383 bp in length
 * 81341 84448: gap of 3572 bp in length
 * 84449 86397: gap of unknown length
 * 86398 86496: contig of 4947 bp in length
 * 86497 91224: gap of unknown length
 * 91225 91324: contig of 4728 bp in length
 * 91325 97769: gap of 6445 bp in length
 * 97770 97869: contig of 478 bp in length
 * 97870 105670: gap of unknown length
 * 105671 105770: contig of 7801 bp in length
 * 105771 111368: gap of unknown length
 * 111369 111468: contig of 5598 bp in length
 * 111469 116880: gap of unknown length
 * 116881 116780: contig of 5212 bp in length
 * 116781 123936: gap of unknown length
 * 123937 123936: contig of 7156 bp in length.
 Location/Qualifiers
 1..123936
 /organism="Rattus norvegicus"
 /db_xref="taxon:10116"
 /clone="CH230-282D16"

AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (24-AUG-1999) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 4 (bases 1 to 168651)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (27-JUN-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 5 (bases 1 to 168651)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (08-NOV-2000) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 COMMENT On Jun 27, 2000 this sequence version replaced gi:7630880.

 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: saplens@wustl.wustl.edu

 Summary Statistics

 Center project name: H_NH0179G23

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality > 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
 Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
 The RPCT-11 human PAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Frengen, E., Teleno, M., Catanesse, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)
 VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:
 The clone sequenced to the left is RP11-547F18; the clone sequenced to the right is RP11-444B4, 200 base pair overlap. Actual end of this clone is at base position 18711 of RP11-444B4.

FEATURES
 source
 1..168651
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="2"
 /map="2"
 /clone="RP11-179G23"
 /clone_1lb="RPCT-11"
 13..437
 /rpt_family="L2"
 repeat_region
 616..683
 /rpt_family="L2"
 /rpt_family="L2"
 1615..1815
 repeat_region
 /rpt_family="MIR"

repeat_region 2577..2748
 /rpt_family="MIR"
 repeat_region 2954..3144
 /rpt_family="MER1_type"
 repeat_region 3149..3465
 /rpt_family="L1"
 repeat_region 3510..3567
 /rpt_family="MER1_type"
 repeat_region 3569..3760
 /rpt_family="MER1_type"
 repeat_region 3827..4198
 /rpt_family="MER1_type"
 repeat_region 5656..6129
 /rpt_family="MALR"
 repeat_region 6205..6316
 /rpt_family="L2"
 repeat_region 6428..6686
 /rpt_family="L1"
 repeat_region 7247..7437
 /rpt_family="MER53"
 repeat_region 7819..7868
 /rpt_family="CR1"
 repeat_region 9304..9469
 /rpt_family="ALU"
 repeat_region 10317..10439
 /rpt_family="MIR"
 repeat_region 12219..12455
 /rpt_family="ALU"
 repeat_region 12467..13537
 /rpt_family="L1"
 repeat_region 13553..14453
 /rpt_family="L1"
 repeat_region 14853..15123
 /rpt_family="MER2_type"
 repeat_region 16209..16345
 /rpt_family="MER2_type"
 repeat_region 16209..16345
 /rpt_family="MER2_type"
 repeat_region 16210..16345
 /rpt_family="MER2_type"
 repeat_region 16513..16676
 /rpt_family="ERVL"
 repeat_region 16970..17198
 /rpt_family="L1"
 repeat_region 17199..17548
 /rpt_family="MER1_type"
 repeat_region 17575..18261
 /rpt_family="L1"
 repeat_region 19409..19531
 /rpt_family="L1"
 repeat_region 19540..19659
 /rpt_family="ALU"
 repeat_region 25486..25822
 /rpt_family="ALU"
 repeat_region 26074..26366
 /rpt_family="ALU"
 repeat_region 26384..26542
 /rpt_family="MIR"
 repeat_region 26569..27058
 /rpt_family="L2"
 repeat_region 27140..27647
 /rpt_family="L2"
 repeat_region 30824..31150
 /rpt_family="MALR"
 repeat_region 32262..32525
 /rpt_family="CR1"
 repeat_region 32568..32770
 /rpt_family="MER1_type"
 repeat_region 32775..33032
 /rpt_family="CR1"
 repeat_region 33172..33418
 /rpt_family="MIR"
 repeat_region 34930..35230


```

repeat_region      /rpt_family="Alu"
                    35302..35555
repeat_region      /rpt_family="L2"
                    35975..36081
repeat_region      /rpt_family="L2"
                    36364..36515
repeat_region      /rpt_family="Alu"
                    37266..37340
repeat_region      /rpt_family="L2"
                    37400..37440
repeat_region      /rpt_family="L2"
                    37615..37715
repeat_region      /rpt_family="MIR"
                    38006..38177
repeat_region      /rpt_family="MERL_type"
                    43171..43204
repeat_region      /rpt_family="U2"
                    44777..44926
repeat_region      /rpt_family="MIR"
                    45905..46110
repeat_region      /rpt_family="MIR"
                    48479..48638
repeat_region      /rpt_family="MERL_type"
                    48914..49033
repeat_region      /rpt_family="MIR"
                    49069..49337
repeat_region      /rpt_family="MERL_type"
                    49340..49621
repeat_region      /rpt_family="Alu"
                    50377..50498
repeat_region      /rpt_family="MIR"
                    56393..56545
repeat_region      /rpt_family="MIR"
                    58109..58420

```

Query Match 27.0%; Score 1157.6; DB 9; Length 168651;
 Best Local Similarity 70.7%; Pred. No. 1.6e-226;
 Matches 2010; Conservative 0; Mismatches 449; Indels 383; Gaps 20;

```

QY 834 AGCGGAGCAGTGGCTGGCAGCCCTGGAAGGAGGCTGAGAGAA-CCCAAGGTATT 892
DB 74820 AGCTGGAGCTCAGAGGGCTCACCCTGCATGGGCTGAGAGAGGCGCGAGTCTACT 74879
QY 893 GCAGGATGATGCTGTGGAAGAAACCAAGGGGCGCGGAGTCTGAGAGGTGAGAG 952
DB 74880 GCAGTATGATGTATGGAAGAAACTAGGGCTCAAGAGAGTGTGCAAGGTGAGAGAG 74939
QY 953 GCCAGCCCTCTGTGTGAGCAGCCCACTGACCCGCTATCCCACTGTGCTACCAAG 1012
DB 74940 GCCAGCCCTCTGTGTGAGCAGCCCACTGACCCGCTATCCCACTGTGCTACCAAG 74999
QY 1013 CCGAGGCGCGTGGGTCGATGCTGGGGAGCAAGATCCCAAGAGAGCGAGATGAG 1072
DB 75000 CCGAGGCGCTGTGGGAGCTGAGGGCTGGGAGCAAGTA---CATCAAGAGGTAGAGATGAA 75056
QY 1073 CCAGAGTACAGAGAGCGCGGCTTTGGCATTGGGAGGCTGCTGGGGAAACTGCGG 1132
DB 75057 CCAGAGTACAGAGAGCACTGGGGCTTTGGCAATGGGAGGCTGCTGAGAGAACTCAG 75116
QY 1133 GGCCTTCTGTGTGGCCAGGCGGCTATGTCTTGTGTGATGACGGGCGGAGCCGAGCA 1192
DB 75117 GGCCTTCTGTGTGT-----GGCGGCTATGTCTTGTGTGATGACGAGCCGAGCA 75171
QY 1193 GCTGAAGGCAACCGCTGGGTGATGTGTGGAGAGAGGCAAAATCTGAGTGTGTGT 1252
DB 75172 GCTGAAGGCAACCACTGTGTGATGTGTGTGGAGAG-6GTGATTTCTAGTGTGTGTGTA 75230
QY 1253 GAGAAGCTGATGCGCTGAGCTCTTTTGGAGTGTGCTTCCACAGGCGCAGS-----TACA 1307
DB 75231 GAGAAGCTGATGACACGAGCTCTTTTGGAGTGTGTGCTTCCACAGGCGCAGCTATATATACA 75290
QY 1308 ACAAGCAGCCCATGTACCGCAAGCCATCTACAGAGTCTTCCAGGTGGCCAGCAGCGCG 1367
DB 1308 ACAAGCAGCCCATGTACCGCAAGCCATCTACAGAGTCTTCCAGGTGGCCAGCAGCGCG 1367

```

```

DB 75291 ACAAGCAGCCCATGTACCGCAAGCCATCTACAGAGTCTTCCAGGTGGCCAGCAG-7GTA 75349
QY 1368 CGGGGAAGCTGTTCGCGGTGTGCACAGCAGCATGTGAGTGTGACACTGCCAAGCCGTGG 1427
DB 75350 CAGGGAGCTGTTCGCGGTGTGCACAGCAGCATGTGAGTGTGACACTGCCAAGCATACAG 75409
QY 1428 AGGTGCAAGAACCCCATGATTTGATGAGGCGCGTGGGGCTTCCAGCCTTCTGGCCCTA 1487
DB 75410 AGGTGCAAGAACCAAGCATATTTGAGATGAGACCTTGGGGGCTTCCAGCCTTCTGGCCCA 75469
QY 1488 AGGGCTTGGAGCCACCAAGAGAGAGAGATCCCTACAAAGATGTACAGGACATGT 1547
DB 75470 AGAGACTGGAGCCCAAGAGAGAGAGAGATCCCAAGAGATTTATGACAGCTGA 75529
QY 1548 GGGTGAACCTGAGGCGCTGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 1607
DB 75530 GGGTGAACCTGAGGCGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 75589
QY 1608 GCACAGCGGAGAGCCCA----- 1625
DB 75590 GCACAGCTAGAGAGCTCAACATGAGGAGCAGCAGCAGATGCTCAACCTGTATCCAGCA 75649
QY 1626 ----- 1625
DB 75650 CTTTGGAGGCTGAGGCGCGGCTGATCAGAGCTGAGAGATGAGAGCTGACATGTGAC 75709
QY 1626 ----- 1625
DB 75710 GCCAATATGTGAAGAACCCCTCTCTACTAAAAAGTACCAAAATAGCTGACATGTGGCA 75769
QY 1626 ----- 1625
DB 75770 TGTGCTGTATATCCCACTACTGCGGAGGCTGAGGAGAGATCACTTACACAGGAG 75829
QY 1626 ----- 1625
DB 75830 TCGAGGTTACAGTACGAGCCAGATGACACCATTTGACACGCTGTGACAGAGCAAGA 75889
QY 1626 ----- 1625
DB 75890 CTCTGTTCAAAAAAATAAAAAAAGCTGACATCAAGGAGATTTGATGA-AC 75947
QY 1652 ACAAGAGAGGCGGTGAGTACGAGGTGCGGAGAGTGGCGAATCATTTGAGCATCTGC 1711
DB 75948 ACAAGAGAGTATGATGATCAAAAGTGAAGAGAGGCGCAAGATCAAGAGCATCTTGT 76007
QY 1712 ATCTCTGTGGAGCTCAATGTATACCTTGAACACCCCTCTTCTGTGAGAGATGTGC 1771
DB 76008 GTCTCTGTGAGAGCTCAATGTATGATGAGAGAGGCGGCTTCAATGCTAGCAATATGA 76067
QY 1772 CAAGAGTGAAGAGCTCTTGTGAGATGTGCGATGAGAGAGAGAGAGAGAGGCTATACAG 1831
DB 76068 CGACATTTGAGAGAGCTCTTGTGAGATGTGCGATGAGAGAGAGAGAGAGGCTATACAG 76127
QY 1832 TCTTACTGCAACATCTCTGTGGGCGGCTGAGAGTGTCTCATGTGCGGAGAAACAACTGC 1891
DB 76128 TTTGCTGCAACATCTCTGTGGGCGGAGAGTGTCTCATGTGCGT---GAACAATGCG 76184
QY 1892 TGCAGGTCTTTTGTGCGAGTGTGTGAGCTCTTGTGTGGGCGGCGGCGCTGCCAGGCA 1951
DB 76185 TGCAGGTCTTTTGTGCGAGTGTGTGAGCTCTTGTGTGGGCGGCGGAGGTTGGCCAGGTG 76244
QY 1952 GGCATTTAGAGAGAGCCCTGAGAGCTGTACATGTGAGGAGCAAGAGGTACTTACGCGCTG 2011
DB 76245 ACCATTTAGAGAGATCCCTGAGAGCTGTACATGTGTGTAACAGAGGAGCATGTGAGAGCT 76304
QY 2012 CTGCGGCGGAGAGAGACTGCGCTCCGCGCTCCAGATGTCTTGTGCTTAATTAACACAG 2071
DB 76305 CTGCGAGTGGGAGATGAGTGCCTGCTCAGCTCCAGATGTCTTGTGCTTAATTAACACAG 76364
QY 2072 CAGGAATTTTACCTCCAAAGGTTTACCACTGTGCCAGCTGAGAGAGAGAG---GCCC 2128
DB 76365 CAGGAATTTTACCTCCGAGAGTCTTACTGCGCTGTGCCAGCTGAGAGAGAGAGAGGCGCC 76424

```

QY	21229	ATTCGGGGGCGTCTCTCTTTGATGGAAATCCCTACAGGGCTCCTGGTGTGAAGACTTGG	2188
Db	76425	ATCCAGGCGCTCTCTTTTGGGGGATGATGATGAGGCGCTGGGCTGTAATGGCTTGG	76484
QY	2189	GGCATTCGATGGAGCCGCTCATTTGGCTGGGAGGTGTGAGGACTCTCATCAAGGTGGCC	2248
Db	76485	GGCATTTAGTAAATCACTACA-TGTCITCAGAGGTGTGTGAGGCGCCATCATTTGGGGC	76543
QY	2249	ATGGGTGGGCGACCGAGGGGAAAGTCAATGTACGTGGGGGAGCGCGAGCGTCAACAGAG	2308
Db	76544	ATGCTGTGGCACCGAGGGAGATCATGCAATCAAGGAAATACCGAATCATATACAGAA	76603
QY	2309	CATATCCAGAGATGGGGCCCATTCGATCTGTGTGANTGGGGGCAATCCCTGCAATGACCTG	2368
Db	76604	CACAAACAGAGATGGGGCCCATTTCT-----CTGCATATGACCTC	76641
QY	2369	TCCATCTGCAACCTGCTGCCAAGGCGCTCTACAGGCGCATGTGGCGGCTCTTCTTTAG	2428
Db	76642	TCCATTTGCAACCTGCGCCAGAGGAACTCTCAAGAGGCACTGGCAGCAATTCCTTGGAG	76701
QY	2429	TTCTACCGCGCTCTGCATGATGCGCGGCCCAAGAGGAGATGATCCGCCCTTCTTCTGG	2488
Db	76702	TTTTTCTCACT-TGTGCATCTATGCAAGCCCAAGGAGGAGATGATGTC-ACCTTCTTGTGG	76759
QY	2489	CTCTTTGGAAATGTGGTGGCCATGGGCGCTTAACTGACAAAGAGGACATCTGCGATTTCTC	2548
Db	76760	CTCTTTGGAGTGTGGTGGCCATGGGCGCACTAGTCAAAAAGAGACATCTCAAAATTTTTT	76819
QY	2549	GAGTCCAAACCTGTGATGATTTGATGCGCAAGAAAGTGTACGTGCACAGGCGCCGCTAC	2608
Db	76820	TATTTGAATTTCCATGATGATTAATATGCCAAGAAAGTGTACGTGCACACCAAGAACTACTAC	76879
QY	2609	TTCTGGGGTAACTCTCCGGTATGAACAGGCGGTTGGCATCGATCGAATGATTAAGCTG	2668
Db	76880	ATCTGTGGTAACTCTCCAGTAAAGACAGGCGCATTTGGCATCTCATATGATG-TAAAGCTG	76938
QY	2669	GAGCTGCAGAGATGCTGAGGACATGCGAGAGTATAGCCAGTTCCACAAAGTGAAGACAT	2728
Db	76939	GAGCTGCAGAGATGCTGAGGACATGCGAGAGTATAGCCAGTTCCACAAAGTGAAGACAT	76998
QY	2729	ACTAGAGGTCAAACTCCATTAAGGAGGGCAAAAGCACAGATTTTCCGTCTTCATGAT	2788
Db	76999	ACTACAGGTCAAACTCCATTAAGGAGGGCAAAAGCACAGCGTCTCCGTCTTCATGAT	77058
QY	2789	GAGAAAGAGACATCTTATGTGTGACTGCAATTCGAAATGGAAGATTTGGTTTCCAGTCCAC	2848
Db	77059	GAGAAAGG-----CACCTTAGGTGATTTGAATTTGAAGAGGATTTTGGACTCTCTGCTCAT	77114
QY	2849	TATACTGACGTCTCCAAATGAGCGCGTTGGGAGGCGACAGACTGTGGCCGCTCATGG	2908
Db	77115	TATACTGACATCTCCAAATGAGCATTTGGGAGGCGACAGACTGTGGCGTGAATGG	77174
QY	2909	AGCGTGCAGTATCCGCACTCTTTCGCTGCGCGTGAAGAGATTTTGGCTGTGTGTA	2968
Db	77175	AGCATGAGCGTCACTGTGCAATTTCTTCCCTCCACGTAAAGGGGTATTTGCTGTGTGTA	77234
QY	2969	GGGACATGGGGCAAACTGAGGTAGCGACACAAAGTTTAAACAAC-----AAACAA	3020
Db	77235	GGGACATGGGGCAAACTGAGGTAGTACACAAAGTTTAAACAACATAAAAAA	77294
QY	3021	AAACACAAAACATAATAAAACACCAAGACATGAGATGAGACAAATACAGACCCAG	3080
Db	77295	CAAAACAAAACATAATAAAACATTAACATGTAAAGAAAGAGAGAAATACAGACCCAG	77354
QY	3081	AAAG-AAAAGAGATTTTAAACAAAACACAGAGGGGAAATCCGAGAGGCTTTGCC	3139
Db	77355	AAAGAGAAAAAGATTTCAAAACAAAACATACAGAGCGAAATTAAGCGAGGCTTGCC	77414
QY	3140	TTTCGAAAAAGGTGTGACATATCTCTGATTTTTCAAATGATCTCTCACTCATTTA	3199
Db	77415	TTTCGAAAAAGGTGTGGA---CATCTCTGATTTTTCAAATGATTTTCTTCACTCATTTA	77471

[illegible]

REFERENCE
AUTHORS
TITLE
JOURNAL

of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 176697)
Morley, K.C.

Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

On Jul 12, 2002 this sequence version replaced gi:20303171.

Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: GRAP

Center clone name: CH230-28122

Summary Statistics

Sequencing vector: Plasmid

Chemistry: Dye-terminator Big Dye: 100% of reads

Assembly program: Phrap: version 0.990329

Consensus quality: 121959 bases at least Q40

Consensus quality: 126371 bases at least Q30

Consensus quality: 130428 bases at least Q20

NOTE: Estimated insert size may differ from sequence length.

(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

NOTE: This is a 'working draft' sequence. It currently

consists of 65 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1 1415: contig of 1415 bp in length

1416 1515: gap of unknown length

1516 2965: contig of 1450 bp in length

2966 3065: gap of unknown length

3066 4228: contig of 1163 bp in length

4229 4328: gap of unknown length

4329 5385: contig of 1057 bp in length

5386 5485: gap of unknown length

5486 6670: contig of 1185 bp in length

6671 6770: gap of unknown length

6771 7813: contig of 1043 bp in length

7814 7913: gap of unknown length

7914 9673: contig of 1760 bp in length

9674 9773: gap of unknown length

9774 11471: contig of 1697 bp in length

11471 11570: gap of unknown length

11571 13162: contig of 1592 bp in length

13163 13262: gap of unknown length

13263 14504: contig of 1242 bp in length

14505 14604: gap of unknown length

14605 15974: contig of 1370 bp in length

15975 16074: gap of unknown length

16075 17396: contig of 1322 bp in length

17397 17496: gap of unknown length

17497 19323: contig of 1827 bp in length

19324 19423: gap of unknown length

19424 21041: contig of 1618 bp in length

21042 21141: gap of unknown length

21142 22419: contig of 1278 bp in length

22420 22519: gap of unknown length

22520 23942: contig of 1423 bp in length

23943 24042: gap of unknown length

24043 25333: contig of 1291 bp in length

25334 25433: gap of unknown length

25434 26530: contig of 1097 bp in length

26531 26630: gap of unknown length

26631 28270: contig of 1640 bp in length

28271 28370: gap of unknown length

28371 29970: contig of 1600 bp in length

29971 30070: gap of unknown length

30071 32265: contig of 2195 bp in length

32266 32365: gap of unknown length

32366 34179: contig of 1814 bp in length

34180 34279: gap of unknown length

34280 36392: contig of 2113 bp in length

36393 36492: gap of unknown length

36493 38552: contig of 2060 bp in length

38553 38652: gap of unknown length

38653 40932: contig of 2280 bp in length

40933 41033: gap of unknown length

41033 43021: contig of 1989 bp in length

43022 43121: gap of unknown length

43121 44977: contig of 1856 bp in length

44978 45077: gap of unknown length

45078 47358: contig of 2281 bp in length

47359 47459: gap of unknown length

47460 49554: contig of 2096 bp in length

49555 49654: gap of unknown length

49655 51877: contig of 2222 bp in length

51878 51976: gap of unknown length

51977 53794: contig of 1818 bp in length

53795 53894: gap of unknown length

53895 55634: contig of 1740 bp in length

55635 55734: gap of unknown length

55735 57445: contig of 1711 bp in length

57446 57545: gap of unknown length

57546 60210: contig of 2665 bp in length

60211 60310: gap of unknown length

60311 62174: contig of 1864 bp in length

62175 62274: gap of unknown length

62275 64285: contig of 2011 bp in length

64286 64385: gap of unknown length

64386 66764: contig of 2379 bp in length

66765 66864: gap of unknown length

66865 69398: contig of 2534 bp in length

69399 69498: gap of unknown length

69499 71451: contig of 1953 bp in length

71452 71551: gap of unknown length

71552 74405: contig of 2854 bp in length

74406 74505: gap of unknown length

74506 76586: contig of 2081 bp in length

76587 76686: gap of unknown length

76687 80397: contig of 3711 bp in length

80398 80497: gap of unknown length

80498 82299: contig of 1802 bp in length

82300 82399: gap of unknown length

82400 84796: contig of 2397 bp in length

84797 84896: gap of unknown length

84897 86981: contig of 2085 bp in length

86982 87081: gap of unknown length

87082 89794: contig of 2713 bp in length

89795 89894: gap of unknown length

89895 91696: contig of 1802 bp in length

91697 91796: gap of unknown length

91797 94458: contig of 2662 bp in length

94459 94558: gap of unknown length

94559 96352: contig of 3794 bp in length

96353 98452: gap of unknown length

98453 101134: contig of 2682 bp in length

101135 101234: gap of unknown length

101235 104588: contig of 3354 bp in length

104589 104688: gap of unknown length

104689 107761: contig of 3073 bp in length

107762 107861: gap of unknown length

111340 111440: contig of 3479 bp in length

111441 115374: contig of 3934 bp in length

Query Match 25.1%; Score 1078.4; DB 2; Length 176697;
Best Local Similarity 79.1%; Pred. No. 2.9e-210;
Matches 1259; Conservative 0; Mismatches 306; Indels 27; Gaps 3;

Oy	1265	TSAGTCTGTTTGAGAGTGCGTTCCACACAGGCCAGCTACCAACAAGACGCCATGTACGCGA	1328
Db	24084	TSAGTCTTTTGAAGATATAAACAATCTCTACGTGAAGTATTTTANNNNNNNNNNNNNNNNNN	24025
Oy	1329	AAGCCATCTACGAGCTCTCGACAGTGGCAGCAGCGCGCGGGAAGCTGTTCGCCGTGT	1388
Db	24024	NN	23865
Oy	1389	GCCACGACAGCGATGAGAGTGAACACTGCCAACGCCCTGAGAGTCAGAACACCCATAGA	1448
Db	23964	NN	23906
Oy	1449	TTGAATGGGCGCTGGGGGGCTTCCACCTCTTGTGGCCCTTAAGGGCTTGAGCCACGAGAAG	1508
Db	23905	TTGAGTGGGCTCTTGGCGGCTTCCACCTCTGTGCTCCCAAGGGCTGTATACCCACCTAAG	23846
Oy	1509	AAGACAACAATCCCTTACAAAGAAGTGTACACGACATGTGGTGTGCAACTGAGAGCAGCTG	1568
Db	23845	ACGACAACAATCTTTACAAAGAATTTTACACGACATGTGGTGTGCAACTCTGAGAGCAGCTG	23786
Oy	1569	CCTAGCGACACCTCCACGACCCCAAAAACCCCGGGAAGAGCACAGCGAAGACCCCAAG	1628
Db	23785	CTTATTTCCTCCCTCCACGACCCCAAAAACCCCAAAAAGAC-----AACCTAAG	23734
Oy	1629	TCAGAGAGATTTATTTGATGAGCGCAACAAGAGCGGCTGTGTACAGATGCGGACAGAT	1688
Db	23733	TCAGAGAGATATTATTTGATGAGCAACGCAAGAATTAACGACTATATATAGAGTCCGCCAATGT	23674
Oy	1689	GGCGAGACATTTAGAGACATCTGCATCTCTGTGGAGCCCTCAATGTATACCTTGGAAACCC	1748
Db	23673	GCGAAGACATTTAGAGACATTTGTATCTCATGTGGAGAACTCCATATGTACCTCTGGACAC	23614
Oy	1749	CCCTCTCGTTGGAGAAATGTGCAAACTGCTTAACTGCTTCTGTGATGAGTGCCTAAC	1808
Db	23613	CACCTCATTTGTGTGATGTGTGCGACAACTGTAAATCTGTTCTTGTGATGTGCTTAC	23554
Oy	1809	AGCAGTACGACAGCGGTACCACTACCTACTGTGACCATCTGCTGTGGGGGGCGTGAAGTGC	1868
Db	23553	AATACGATGACGATGTGTACCACTCTTGTATCCATCTGCTTGTGGGGGGCGGAAGTGC	23494
Oy	1869	TCATGTGGGGAACAACAACGCTGCGAGAGTGTCTTTGGTGAGTGTGTGTGAGACCTCTTGG	1928
Db	23493	TCATGTGTGGAACAACAACGCGTGGCAGGTGCTTTGTGTGTGAGTGTGTGTGATCTTGG	23434
Oy	1929	TGGGGCGGGGGGCTGCCAGGACAGCCATTAAAGAAAGCCCGGAGACGCTACATGTGGG	1988
Db	23433	TGGGGCGAGGGGCTGCCAAGACAGCCATTAAAGAAAGACCCCTGTATACGTCATCATGTG--	23376
Oy	1989	GGCACAAAGGTACCTACGCGGCTGTGGGGGGCGGAGAGACGTGCGCCCTCCCGGCTCCACA	2048
Db	23375	-----TGAGCTGTGGGGAAGGAGAGACTGGGCTTCTTCGACTCCACA	23332
Oy	2049	TGTTCTTCGCTAATTAACACAGACAGCAATTTGACCTCCCAAAAGTTTACCACCTGTGC	2108
Db	23331	TGTTCTTTTCCCAATTAACACACACAGCAATTTGACCCCCCAAGTTTACCACCTGTGC	23272
Oy	2109	CACCTTAGAAGAGAAAGCCATCCGCGGTGTGTCTCTCTTTGATGGAATGCGTACAGAGGC	2168
Db	23271	CACCTTAGAAGAGAAACCCATCCACGCTGTATCTCTATTGTATGGATGTGTCTACAGGGC	23212
Oy	2169	TCTGTGTGTGAAGACTTGGGCATTCAGTGTGACCGCTACATTTGCTCGAGAGTGTGT	2228
Db	23211	TCTGTGTGTGAAGACTTGGGCATTCAGATGTGACCGCTACATTTGCTCGAGAGTGTGT	23152
Oy	2229	AGGACTTCATCAACAGTGGGCATGTGTGCGACACAGGGGAAGCATCTAGCTGTGGGAGCG	2288
Db	23151	AGGACTTCATCAACCTGGGCATTAATGTGGGACACAGGGGAAGCATCTAGCTGTGGGAGCG	23092
Oy	2289	TTCGACGCTCACAGAAAGCATATTCAGAGAGTGGGGCCATTGCATCTGGTATTTGGG	2348
Db	23091	TTCGACGCTCACAGAAAGCATATTCAGAGAGAGGGGCCCATATCAATCTGGTATTTGGG	23032

OY	2349	GGAGTCCCTGGCAAGACCTCTCCATCGTCAACCCGTCGCGAAGAGGCTCTACGAGGGCA	2408
Db	23031	GCAGTCTTGGCAATGACCTCTCCATCATCAACCCGCGCAAGGACCTTACAAAGGGCA	22972
OY	2409	CTGGCGGGCTCTCTCTTTGAGTTTCTTACCGGCTCTGATGATGCGCGCCCAAGAGGAG	2468
Db	22971	CTGGCAGGCTCTTCTTTGAGTTCTTACCACTCTCTCATGATGATGCGCGCCCAAGAGGAG	22912
OY	2469	ATGATCGGCGCCCTCTCTCTGCGCTCTTTGAGCAATGAGTGAGGCCATGAGGATGAGCAAA	2528
Db	22911	AAGATCGGCGCCCTCTCTCTGCGCTCTTTGAGCAAGTGAGGCCATGAGGATGAGCAAA	22852
OY	2529	GGGACATCCCGCGAATTTCTCGAGCTCCCAACCCGTATATGATGATGCCAAAGATGCTCAG	2588
Db	22851	GGGACATCCCGCGAATTTCTTAGACTTAAGCTGATGATGATGATGATGATGATGATGATG	22792
OY	2589	CTGCACACAGGAGCCCGGTACTCTCGGAGTAAACCTTCCCGGTATCAACAGCCCTGTGCAT	2648
Db	22791	CTGCACACAGGAGCCCGGTACTCTCGGAGTAAACCTTCCCGGTATGATGATGATGATGATG	22732
OY	2649	CCACATGTAAGTAATTAACCTGAGCTGAGAGTGTCTGAGACATGCGCAGATAGCCAACT	2708
Db	22731	CCACATGTAAGTAATTAACCTGAGCTGAGAGTGTCTGAGACATGCGCAGACACCCAGCT	22672
OY	2709	TCAGCAAGTGAAGACCACTTACTACGAGGTCAAACTCCATTAAGCAGGGCAAGACCAAC	2768
Db	22671	TCAGCAAGTGAAGACCACTTACTACGAGGTCAAACTCCATTAAGCAGGGCAAGACCAAC	22612
OY	2769	ATTTTCCTGTCTTCATGATGATGAGAAAGACACTTTATGTGTGCACTGAATGGAAGGG	2828
Db	22611	ATTTTCCTGTCTTCATGATGATGAGAAAGACACTTTCTGTGTGCACTGAATGGAAGAC	22552
OY	2829	TATTTGGTTTCCCACTCCACTATACAGCTC	2860
Db	22551	TGTTTGCTTACCCCTGCTCTACTACAGAGCC	22520
RESULT	11		
AF176228			
LOCUS	AF176228	4267 bp	mRNA linear PRI 26-OCT-1999
DEFINITION	Homo sapiens DNA cytosine-5 methyltransferase 3B (DNMT3B) mRNA,		
ACCESSION	AF176228		
VERSION	AF176228		
KEYWORDS	AF176228.1	GI:6118091	
SOURCE			
ORGANISM	Homo sapiens.		
	Homo sapiens.		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
REFERENCE	1 (bases 1 to 4267)		
AUTHORS	Xu,G.-L., Bestor,T.H., Bourc'h,H.S.D., Hsieh,C.-L., Tommerup,N.,		
	Bugge,M., Hullen,M., Ou,X., Russo,J.J. and Viegas-Pequignot,E.		
	Chromosome instability and immunodeficiency syndrome caused by		
	mutations in a DNA methyltransferase gene		
	Nature (1999) in press		
TITLE	2 (bases 1 to 4267)		
	Xu,G.-L. and Bestor,T.H.		
JOURNAL	Direct Substitution		
REFERENCE	Submitted (06-AUG-1999) Genetics and Development, Columbia		
AUTHORS	University, 701 West 168 St., New York, NY 10032, USA		
TITLE	Location/Qualifiers		
JOURNAL	1. 4267		
FEATURES	/organism="Homo sapiens"		
SOURCE	/db_xref="taxon:9606"		
	/chromosome="20"		
	/map="20q11-q13"		
	1. 4267		
gene	/gene="DNMT3B"		
CDS	265..2802		
	/gene="DNMT3B"		
	/note="mutated in human chromosome instability and		
	immunodeficiency disease known as ICF syndrome; contains		
	alternative 5' exon 1B"		

```

/codon_start=1
/product="DNA cytosine-5 methyltransferase 3B"
/protein_id="AA04015.1"
/db_xref="gi:6118092"
/translation="MEPSPEPSLESKGDTRIHNGEDNAGREDSTLVNAGSDSS
DSPLELAITRPIELRGRSSRLSKREVSLISTYQDLTGDEDEDSDSTPYMPL
FRFRTRESBPVATRRNNNSVSSNERHPRSRGROGRHDESEVEPATSLR
RATASAGTPMPSPSXYLTIDLTDEDTPDTPDSTPOSSTPYARLADSOOGMESPOE
ADSGDSESEYODKEFGIDLVNGKIKGSMWPAMVSNKATKIRAMSGMBWOMF
GDKESEVSADKIVAGLGLFSOHENLATENLVSYRKMYHLEKARACKTPESSBG
DSLEDLAKPLERAHGFKPTGIBGLKPNTPENKTRRRIRADDSATSDCPARKK
TNCNNKDKGDEDSREQASDVANNKSLSDGLSCGRNPVSEFPLEGGICQCPC
RDRELELYMDDEDSQSYCYCCCEGRELISCNTSCRCFVCLVLTGTAADA
KLOEPMSCYMLQPRCHGVLRRKDMVRAFTSPDLEEARKEPLPAIPARRRP
IRVSLFDGIATGLVLEKGIKYGKVASVECESTAVGVKHEGNIKYNDVRNT
KNNEEMGPDIIVIGSPNDLSNVNPARGLYRGTRLPPEPHILNLSRKPGRDR
PFRMPFNVYAMKVGDKRDISRFLPCNPMDITKVSNAHARLTFWNLPGMNPVIA
SKNDLELDCLLEYNRIAKLKVQITTTKSTIKQGNLFTPVVNMNKEEDVLMCTELE
RIFFEPHYHYDVSNMGRGARKILGRMSVPVIRHLPFLKDYFACE"
BASE COUNT      1069 a      1098 c      1147 g      953 t
ORIGIN

```

```

Query Match      18.4% Score 791.8; DB 9; Length 4267;
Best Local Similarity 64.0%; Pred. No. 1.1e-151;
Matches 1248; Conservative 0; Mismatches 667; Indels 36; Gaps 2;

```

```

QY 1017 AGCCGTTGGGGTCCGATGCTGGGAGCAAAATGCCAACAGAGCGCATGACGACG 1076
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 887 AGCAGGGGGGCGATGGAGTCCCGCGAGTGGAGCGACAGTGGAGATGACAGCTTCAG 946
QY 1077 AGTACGAGAGCGCGCGGGGCTTTTGCAATTGGGAGCTGGTGGGGGAAATCGCGGCT 1136
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 947 AGTATCAGAGATGGGAGAGTTTGGATAGGGGACCTCGTGTGGGAAAGATTCAGAGGCT 1006
QY 1137 TCTCTGTGTGCGCAGCGCCGATTTGTCTGTGTGGATGACGGCGGAGCCGAGCAGCTG 1196
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1007 TCTCTGTGTGCGCGCCATGCTGTGTGTGGAAGGCGCCCTCCAGGAGCAGCCTATGT 1066
QY 1197 AAGCAGCCCGCTGGGTGATGTGTGAGAGCGCAAAATTCAGTGGTGTGTGTGA 1256
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1067 CTGGCATATGCGGTGGGTCCAGTGTGTTTGGCGATGCGCAAGTCTCCGAGGTCTCTGACAGCA 1126
QY 1257 AGCTGATGCGCGTGAAGCTGTTTTCAGTGTGCTTCACACGAGCCACGTACAAACAAGCAGC 1316
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1127 AACTGTGTGCGACTGGGGCTGTTTCAAGCAGCACTTTAATTGGCCACCTTCAATTAAGCTCG 1186
QY 1317 CCATGTACCGCAAGCCATCTACAGAGTCTCTGCAAGTGGCCAGCAGCCCGCGGGGAAGC 1376
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1187 TCTCTATCGAAAGCCATGTACATGCTCTGGAGAAAGCTAGAGGTGCGACCTGGCAAGA 1246
QY 1377 TGTTCGCGGTGTGCGACAGCAGATGAGTGAACCTGCCAAGCGCGTGGAGGTGCGA 1436
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1247 CCTTCCC-----CAACAGCCCTGGAGACATCATTTGAGAGACACAGC 1285
QY 1437 ACAAGCCCATGATTGAATGGCCCTGCGGGGCTTCACAGCTTCTGCGCCCTAAGGGCCTGG 1496
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1286 TGAAGCCCATGTTGAGTGGGCCCAAGCGGGGCTTCACAGCCCATGCGAGTTCAGAGGCGCTCA 1345
QY 1497 AGCCACCAAGAGAGAGATCCCTACAAAGAGTGTACAGGACATGTGGGTGGAAC 1556
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1346 AACC-----CAACAACACGCAACAGAGAACAGACTGTGAAGAGCGACAG 1390
QY 1557 CTGAGGCGAGTCTAGCGACACCTCCACGCAAAAGAGCCCGGAGAGAGCAGAGG 1616
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1391 CTGAGGACTCAGCCACCTCTACTACTGCCCCGACCCCAAGCGGCTCAAGCAAAATTGCT 1450
QY 1617 AGAAGCCCAAGTCAAGAGATTATGTATGAGCCGACAAAGAGAGCGGCTGTGATGAGG 1676
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1451 AFAACAAGCGCAAGAGCCGAGGAGATGAAGATCAGAGCCGAGAAACAATGGCTTCAGATG 1510
QY 1677 TGGCGAGAGTGGCGGACATTTGAGACATCTGCATCTCTGTGGAGAGCTCAATGTTTA 1736
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1511 TTGCCAACACAAAGAGCAGCTTGGAAGATGGCTGTGTTGTTGTGGCAGAGAAACCCCG 1570

```

```

QY 1737 CCTTGAAACACCCCTCTTCTGTGGAGGATGTGCCAAAACTGCAAGAACTGCTTCTGG 1796
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1571 TGTCTCTCCACCTCTCTTTTAGGGGGGGGCTGTGTACAGACTGCCGGAGATGCTTCTTG 1630
QY 1797 AGTGTGGTACCACTAGCAGACAGCAGCGCTACACAGTCTCTACTGACCATCTGCTGTGGG 1856
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1631 AGCTGTTTATCATGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1690
QY 1857 GCCGTGAGTCTCTATGTGGCGAAACAAACACTCTGAGGTGCTTTTGGCTGGAGTGTG 1916
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1691 GCCAGAGCTGCTCTCTTTCAGCAACAGCACTCTGCGGCTGTGTTCTGTGTGAAGTGC 1750
QY 1917 TGGACCTCTTGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1976
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1751 TGGAGGTCTGTGTGGGACAGCAGCAGCGCGCGGAGCCAGCTTCAGAGAGCCCTGAGAGT 1810
QY 1977 GCTACATGTGGCGGACAGAGGATCTACAGGGCTGTCTGCGGCGGCGGAGAGAGCTGGCCCT 2036
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1811 GTTACATGTGTCTCCCGAGCGGTGTATGCGCTGCTGCGGCGCGGAGAGAGCTGGAACG 1870
QY 2037 CCGGCTTCAGATGTTCTTCTCTATATACACAGCAGAGATTTGACCTTCAAGGTTT 2096
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1871 TGCCTGTGACAGCGCTTCTTACACAGTGCACAGGGGCTTGAATATGAAGCCCAAGCTGT 1930
QY 2097 ACCACCTGTCCAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2156
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1931 ACCCTGCCATTCGCCGACACCCGAGAGCGGCGCCATTCAGTCTGTCTGATTTGATGGCA 1990
QY 2157 TCGCTACAGGAGCTCTGTGTGTGAAGAGACTTGGGCAATTCAGTGGAGCCGCTACATTCGCT 2216
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1991 TCGGACAGAGCTTACTAGTACTCTCAAGAGATTGGGCAATTAAGCTAGAAAGAACGCGCTT 2050
QY 2217 CGAGAGTGTGTGAGACTCTCATCAGCGTGGGCAATGTGTGGCCAGCAGGAGAGATCATGT 2276
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2051 CTGAAGTGTGTGAGAGTCCATGCTGTGTGAACCGTGAACAGCAGAGGAGATATCAAT 2110
QY 2277 ACGTGGGAGCTCGGAGCTGACACAGCAAGCAATTCAGAGAGTGGGCGCCATTCAGTTC 2336
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2111 ACGTGAACAGCTGAGAGCAATCATCAAAAGAAATATTGAAGATGGGCGCCATTTGACT 2170
QY 2337 TGTGATTTGGGGGAGTCCCTGCAATGACCTCTCATGTCACACCTGCTGCAAGGGCC 2396
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2171 TGTGATTTGGGGGAGAGCCCATGCAACGATCTCTCAAAATGTGAATTCACACGAGAAAGGCC 2230
QY 2397 TCTACAGAGGAGTGGCGGCTCTTCTGTAGTCTTACCGGCTCTGTGATATGCGGGC 2456
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2231 TGTATGAGGATACAGGGCGGCTCTTCTGCAATTTTACCACTGCTGAATTAATCAGCGC 2290
QY 2457 CCAAGGAGGAGATGATGCGCCCTTCTGTGCTCTTGAATGTGGTGGCCATGAGGCG 2516
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2291 CCAAGGAGGAGTATGACCGGCGCTTCTGTGATGTTTGAATATGTTTATACCATGAAGG 2350
QY 2517 TTAGTACAGAGGAGATCTCGCGATTCTCGATGCTCAACCTGTGTATGATGATGCA 2576
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2351 TTGGGCAACAGAGGACATCTCAGCGTCTCGAGATGTAATTCAGATGATGATGATGCA 2410
QY 2577 AAGAAGTCTACGCTGACACAGGCGCGCTTCTGTGGGTATACCTTCCGCTATGAGCA 2636
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2411 TCAAAGTCTGTGCTCTACAGGGCGCGATCTTGTGGGAAACCTTACCGGAGTGAACA 2470
QY 2637 GCGGCTGTGATTCACAGCTGTGAATGATTAAGCTGAGAGCTCAGAGAGTGTGTGAGCATGCA 2696
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2471 GCGCGGTATATGATCAAAAGATGTAATTAAGCTGAGAGCTCAGAGATGCTTGGATATCAATA 2530
QY 2697 GGATAGCCCAAGTTCAGCAAGATGAGAGACATTTACTAGAGGTCAAACTCATTAAGCAGG 2756
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2531 GGATAGCCCAAGTTCAGCAAGATGAGAGACATTTACTAGAGGTCAAACTCATTAAGCAGG 2590
QY 2757 GCAAAGACCAAGATTTCTCTGTCTTATGATGATGAGAAAGAGACATCTTATGTGCTGACTG 2816
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2591 GGAAGAAACCAACTTTTCCCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 2650

```

Accession	Sequence	Position
QY	2811 AAATGGAAGGATATTTGGTTTCCAGACACATATACGTCTCCCAATAGAGCGGT	287/6
Db	2651 AACTGGAAGGATCTTTGGCTTCTCTGACACTACACAGAGGTGTCACATGTGGCGTG	2710/0
QY	2877 TGGGAGGACAGACATCGTGGGGCGGTCATGAGAGCGTGTCAATCATGCGCACTCTTG	2936/6
Db	2711 GTGGCCGCCAGAGATGCTGCGGAAGTGTCTGAGAGCGTGTCTATCTGACACCTTTTG	2770/0
QY	2937 CTCCGCTGAAGGAGTATTTTGGCGTGTGTGA	2967
Db	2771 GCCCTGTGAAGGACTACTTGTGATGTGATAA	2801

RESULT 12	LOCUS	DEFINITION	VERSION
AF068627	4135 bp	mus musculus DNA cytosine-5 methyltransferase 3B2 (Dnm3b) mRNA, alternatively spliced, complete cds.	AF068627.2
			GI:6449471

EXTRACTO: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Rodentia: Sotugonathi: Muridae: Murinae: Mus.
REFERENCE
TITLE
AUTHORS
 1 (bases 1 to 4135)
 Okano, M., Xie, S. and Li, E.
Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases
JOURNAL
NO. Natl. Gene. 19 (3), 219-220 (1998)

REFERENCE AUTHORS TITLE JOURNAL	REFERENCE AUTHORS TITLE JOURNAL
2 (bases 1 to 4135) Xie,S., Okano,M. and Li,E. Direct Submission Submitted (28-MAY-1998) CVRC, Mass. Gen. Hospital, 149 13th Street, Charlston, MA 02120, USA	3 (bases 1 to 4135) Okano,M., Chijiwa,T., Sasaki,H. and Li,E.

JOURNAL Submitted (04-NOV-1999) CVRC, Mass. Gen. Hospital, 149 13th Street,
Charlestown, MA 02129, USA
REMARK Sequence update by submitter
COMMENT On Nov 18, 1999 this sequence version replaced gi:3327981.
FEATURES
 source Location/Qualifiers
 1. .4135

```

/db_xref="taxon:10990"
/chromosome="2"
/map="84.0 CM"
1. .4135
/gene="Dnmt3b"
misc_feature
1. .4135

```

/note="similar to EST sequences deposited in GenBank
 Accession Numbers AA116694, AA119979, AA177277, AA210568,
 AA407106, and AA575617"
 269. 2788
 CDS

/note="alternatively spliced product; contains Cys-rich
 region; C-terminal region is similar to corresponding
 region of DNA cytosine-5 methyltransferase 3A"
 /codon_start=1
 /product="DNA cytosine-5 methyltransferase 3B2"
 /protein_id="AAC40179.2"
 /db_xref="GI:6449472"
 /translation="MKDSDHLNEPQAGYECCITNGNFSNOSDTRKAPSPYLE
 AITQPCPEPAGCRASSRSLSKRPYSLLIAYTDMDCGDDRDDEIDGKNSDILMK
 LITEETDTRTRESPPAIRTRHSGTSLERKQASRTLRGQRGRIHVEEIPK
 SRRRRSSASTPWSFASVDFMEETVETSTSVSLDSGGQSDMTDVIDEAD
 GSTETQDDKEFGIGDLVMGKIKGFNSKAAVSMKTRSLRQAKPMQKRVWVGDGFE
 SETSAKRLVAGLEISQHPNINATINVTYSRAAMHTLEKRAVAGATCTSSPGSGLSD

BASE COUNT 1043 a 1048 c 938 t

Best local similarity 64.6%; Pred. No. 5,9e-151; Mismatches 1262; Conservative 0; Mismatches 648; Indels 45; Gaps 4.

OY 1016 GAGCCCTGGGCTCGATGCTGTGGGACACAGATGCCACCACAAAGCAGCCGATGCAGACCA 1075
 ++++++ |++| ++ |++| ++++++ |++| ++++++ |++| ++++++ |++| ++++++
Db 875 GATCAGAGGGTGTGATCACACACAGTGTCGAGAGACAGAGATGTGAGACAGACACA 934
 ++++++ |++| ++ |++| ++++++ |++| ++++++ |++| ++++++ |++| ++++++

OY 1136 TTTCCCTGGTGGGCGACGCCGCATTTGCTTCTTGTTGATGACGGGCCGGAGCCGACACT 1135
 |||||
 Db 995 TTTCCTCGTGGGCTGCTCCCTGGGTGTGTCCTGGAAAGCCACCTCCAAAGCAGCAGGGCATG 1054
 |||||
 OY 1136 GAAAGCAGCCCGTGGGTGTCATGTGTTGCGAGACGCGCAATTCATGATGGTGTGTGAAG 1255
 |||||

OY	1256	AAAGTGAATGCCCGTAGAGCGCTTTGGAGAGCGCTTCACACAGGGCAGATCAACACAGG	1315
Db	1115	AAATCGTGGCTGTGGGGGTGTTCAGGCACACATTTAATCTGGCTACTCAATTAAGCG	1174
OY	1316	CCCATGTACCGCCAAACCATCTACGAGGTCTTCGACAGTGGCCAGCAGCGCGGGGGAAG	1375
Db	1175	GTATCTATATGGAAGGCCATGTGACACACTCTGGAGAAAGGACCGAGGTTCAGCTGGCAAG	1234

OY 1376 CTGTTCCCGGTGGCAGACAGGATGAGTGAAGACTGGCAAGGCCGAGAGGTGCAG 1435
||| ||| ||| ||| ||| |||
Db 1235 ACCTTC-----CAGAGTCCTGGAGAGTCACTACGAGGACAG 1273
||| ||| ||| ||| ||| |||
OY 1436 AACAGCCCATGATTGAGCCCTGGGGGCTTCACAGCTTGCCCTTAAGGCGCTG 1495
||||| ||| ||| ||| ||| |||

[illegible]

Db 1388 TCTGCTGCTTCTGAGTCCCCCCACCACCGCCCTAAGACAAATAGCTATGCGCGGGGAAG 1447

Oy 1616 GAGAAGCCCAAGTCTCAGGAGATTATTGATGAGCGCACAAAGAGACCGCGCTGTACGAG 1675

QY	1736	ACCCGTGAAACACCCCTCTTGTTGGAGAAATGTGCAAAACGTCAAGAAACGTCTTG	1795
Db	1553	GTGTCCCTTCACACCCCTCTTTGAGAGGTGGGCTCTGTCAAGATTTGCCGGATGCGTTCTTA	1612
QY	1796	GAGTGTGCGTACCAAGTACGACGACGAGCGGTACACGCTCTTACTGACACATCTGCTGTGGG	1855
Db	1613	GAGTCTTCTTCAATGAATGAGAGAGGGGTATCAAGTCTTCTGACACCGGTGTGCTATAG	1672
QY	1856	GGCCGTGAGGTGCATGTCCGCAAAACAACACTGCTCAGCTGCTTTTTCGCTGGAGGT	1915

Db	1673	GGCGGTGAACGCTGCTGTGTGACGTAAACAAAGCTGCTGCAGATCTCTGTGTGSAATGT	1732
OY	1916	GTGACACCTCTTTGTGGGGCGGGGGCTGCCAGGACCATTTAAGAAAGACCCCTGGAAC	1975
Db	1733	CTGGAGGTGCTGTGGGGCGGCAGGACAGCTTAGAGATGCCAAAGCTGCAGGAACCTGGAGC	1792
OY	1976	TGCTACATGTGGGGGACAAAGGTACTAGGGGCTGTCGGGGCGGCGAAGAGATGGCCC	2035
Db	1793	TGCTATATGTCCCTCCACGCTGTGCATGGGGTCTCCGACCCAGAGAAAGATTGGAAC	1852
OY	2036	TCCGGGCTCCGATGTCTCTTCCGTATATACCAACAC --- AGSAATTTGACCCCTCAAG	2092
Db	1853	ATGGGCTCGAAGACTTCTTACTACTGATCTCTACCTGGAAAGATTTGAGCCACCAAG	1912
OY	2093	GTTTACCACCTGTCCAGCTGAGAAGAGAACCCATCCGGGTGCTGTCTCTTTGAT	2152
Db	1913	TTTGTACCAGCAATTCCTGCAGCCAAAAGAGGCCCATTTAGAGTCTGTCTGTTTGAT	1972
OY	2153	GGAAATCGGTACAGGGCTCTGTGCTGTAAGACTTTGGGCAATTGAGGTGACCCGTACAT	2212
Db	1973	GGAAATGTCAACGGGGGTACTGTGCTCTCAAGAGAGTTGGGTATTAAAGTGAAATATACAT	2032
OY	2213	GCCTCGAGAGTGTGTGAGAGACTTCATCAACGGTGGGCAATGGTGGCGCACCGAGGAAGTC	2272
Db	2033	GCCTCCGAAAGTCTGTGCAGAGTCCATGCTGTGGGAATGTTAAAGCATGAAAGGCGAAGTC	2092
OY	2273	ATGTACGTCGGGGAGCTCCCGAGGCTACACAGAAAGCATATACAGAGTGGGGCCCATTC	2332
Db	2093	AAATATGTCAATGACGTCGGGAAATATCACAAAGAAATATTGAAGTGGGGCCCGTTC	2152
OY	2333	GATCTGGTGATTTGGGGGCGACCTCTGCATATACCTCTTCATCGTCAACCCCTGCTCGAAG	2392
Db	2153	GACTTGGTGATTTGTGGAAGCCCATGTGCATGTCTCTAACTGTCATATCCCTGCCGGAAA	2212
OY	2393	GGCCTCTACAGAGGCGACAGCGCCGCTCTCTTGTAGTTCTACGCCCTCTGCATGATGCG	2452
Db	2213	GGTTTATATGAGGGGCACAGAAAGCTCTTCTTCAGATTTTACCACTTGCTGTAATTTATCC	2272
OY	2453	CGGCCCAAGAGGAGAGATGATGCGCCCTTCTTCTGCGCTCTTTGGAATGTGTGGCCATG	2512
Db	2273	CGCCCAAGAGGGGCGAACCGTCCATTTCTTGGAATGTTGCGAAGTGTGTGGCCATG	2332
OY	2513	GGCCTTACTGCAAGAGGAGACATCTCCGATTTCTCGATCCAAACCCCTGTGATGATGAT	2572
Db	2333	AAAGTGATGATCAAGAAAGACATCTCAAGATTTCTGCGATGTATACCCAGATGATGATGAT	2392
OY	2573	GCCAAAGAGTGTAGCTGCACACAGGGCCCGCATCTCTGGGGTAAACCTTCCCGGATG	2632
Db	2393	GCCTCAAGGTGTGTGCTGCACACAGGGCCCGTACTTCTGGGGTAACTAACCCGGAGTG	2452
OY	2633	AACAGGCGTGTGGATCCACTGTGATATGATTAAGCTGTGAGCGACAGAGTGTCTGGACAT	2692
Db	2453	AACAGGCGCGTGTGCTGCACACAGGGCCCGTACTTCTGGGGTAACTAACCCGGAGTG	2512
OY	2693	GGCAGGATAGCCAAAGTTCAGCAAAAGTACACCATTAATCTACAGAGTCAAACTCATTAAG	2752
Db	2513	AGTGTGACAGCAAAAGTTTAAAGAAAGTGCAGACATTAACACCAAGTGCAGACTCATGAGA	2572
OY	2753	CAGGGCAAGACACGACATTTCTGCTCTCAATGAAATGAGAAAGAGACATTTATGTGTC	2812
Db	2573	CAGGGCAAAACACAGCTTTCTCCGTGTGATCAAGATGCAAGAGACAGCTTTTGTGTGC	2632
OY	2813	ACTCAAAATGAAAGGATTTTGTGTTTCCCATCTCCATATACTGACGATCCCAATGAGC	2872
Db	2633	ACTGAGCTCGAAAGATTTCTGGCTTCCCTGCTCACTACACAGGACGTGTCCAAATGAGC	2692
OY	2873	CGCTTGGCGAGGACAGACAGCTGTGGCCGCTCATGAGAGCGTGCACATCTCCGACCTC	2932
Db	2693	CGCGCGCGCCCTCGAAGAGTCTGGGAGAGTCTCTGGAGTGTACCGGTCATCAGACACCTG	2752
OY	2933	TTGCTCGCGTGAAGAGATATTTTGGGTGTGTGA 2967	

DB	2753	TTTGCCCCCTTGAAGGACTACTTTGGCCGTGAATAA	2787
RESULT 13			
LOCUS	AF151974	4163 bp	mRNA linear ROD 08-JUN-2000
DEFINITION	Mus musculus DNA cytosine-specific methyltransferase isoform 6		
ACCESSION	AF151974		
VERSION	AF151974.1		
KEYWORDS	(Dnmt3b) mRNA, complete cds.		
SOURCE	Mus musculus.		
ORGANISM	Mus musculus.		
REFERENCE AUTHORS	Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Eutelesomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 4163)		
TITLE	Yin,B., Chen,Y.T., Zhu,M., Luo,Y.J., Zhu,N., Xu,S.C., Wu,G.Y. and Shen Y.		
JOURNAL REFERENCE AUTHORS	Cloning of full-length Dnmt3b cDNA and its alternative splicing isoforms in mouse embryonic tissue Unpublished 2 (bases 1 to 4163)		
TITLE	Direct Submission Submitted (17-MAY-1999) Department of Biochemistry and Molecular Biology, Institution of Basic Medical Sciences, Chinese Academy of Medical Sciences, Dong Dan Shan Rao 5, Beijing City 100005,		
FEATURES	P.R.China		
source	Location/Qualifiers 1..4163 /organism="Mus musculus" /strain="KM" /db_xref="taxon:10090" /dev_stage="-8-9 day old embryo" 1..4163 /gene="Dnmt3b" 297..2816 /gene="Dnmt3b" /function="methylates cytosine in DNA" /note="alternatively spliced" /codon_start=1 /product="DNA cytosine-specific methyltransferase isoform 6"		
gene	/protein_id="AAF74520.1" /db_xref="GI:8347131" /translation="MKDGRLNEEGEAGSGYEELIYNGNFSQDSOTKAPSPVRE AICHPVCPTETRGRRSSRSLSKRDESSLATYTDMTDGDDRDEVDGNSDLIMPLKE LTRERKDTRTSESPPAVTRHNSQTSLERORASPRITRGKHGHVOEPVEFPAIR SRRRASADSTPWSPPASVPMEBEVTPKSTPSYVLSDGDGEDMDTTOVADESIST GDSLEYODDCEFGIGIDLAWCKIKGFSWMPAAVMWMAITSRKQAMPGRMWOMFDGKFL SEISADKLVALGIFSOHNLTLPKNLVSRYKAMVHTLEKARVAKGFESSPGSELEGE OLKPMLEAHGCGFKPQTEGLEKPKKKOPEKSRRTTNDSASSPPKRLTKTSYVG KDRGDEESRRMASVTNNKGNLEDRLSCGKANPVSFRHLFEGGLCQSQRDFLEI FYMDEDYOISTCTVCCGRELILCSNTSCRCCEVELEVGVAGTAEDKLDPMS CYNMLPORCHGYLRRRDKMMNRLODFTTDDLEFEFPKLYPAIPAKRPDIVLSL FDGLATGYLVLEKEIGIKVEKYIASVECAESIAGTVHEGOIKYVNVRIKTKNIEEF WGPDELIVGSPENDLSINVNPARGIVEGTGRLLFFEFYHLINTYRPEGDNRPFEEMF ENVMVAMYNDKKDISRLACPNVMIIDAITYSAARVARVFNGLPMRNPVWASKNDKITI ELQCLLESRTIAKKTQVTTTYSKNTROGNOLEPFVYMKGKDYLMTLELERLFEGFP AHYDVSNMGGAOKKLLGRWSVSVPIRHLPALPDYFACE"		
BASE COUNT	. 1032 a 1074 c 1128 g 929 t		
ORIGIN			
Query Match	18.4% Score 788.2 DB 10 Length 4163:		
Best Local Similarity	64.6% Pred. No. 5.9e-151:		
Matches 1262;	Conservative 0; Mismatches 646; Indels 45; Gaps 4;		
DB	1016 GAGCCCGTGGGGTCGATCTGCGGGGACAAGAATGCCACCAAAGCAGCGATGACGAGCCA 1075		
Db	903 GATCAGGAGGGATATGATACACACAGAGTGATGCAGAGACCATATATGAGACAGCACCA 962		
0y	1076 GAGTAGCAGGACGCCCGCTTGGCATTTGGGAGAGCTGCTGTGGGGGAAACTCGCGGGC 1135		

Db	963	GAGTATCGAGGATGATAAACAGTTTGGAAATAGTGACCTGTGTGGGAAAGATCAAGGGC	1022
Oy	1136	TTCTCTCGTGGGGCAGGCCCGCATTTGTGTCTTGGTGTGATACGGGGCCGAGCCACAGAGCT	1195
Db	1023	TTTCTCTCGTGGGCTCCATCGTGTGGTTCCTGGAAAGCCACCTCCAGCCACAGGGCCATG	1082
Oy	1196	GAAAGCACCCTGCGGTCACTGTGTGGAGAGGCCAAATTTCTAGTGGTGTGTGTGAG	1255
Db	1083	CCCGGAATGCGGTGGGTACAGTGGTTTGGATGTGGCAAGTTTCTGAGATCTCTGTGAC	1142
Oy	1256	AAGCTGATGCGCGTGAAGCTCGTTTGGCGTGGCTTCACACAGGCCACGTACACAGCAG	1315
Db	1143	AAACTGGTGGCTGTGGGGTGTTCAGCCACACTTTTAATCTGGGTACCTTCATTAAGCTG	1202
Oy	1316	CCCATGTACCGCAAAACCATCTACGAGGTCTCTAGTGGGGCAGCAGCCCGGGGGAG	1375
Db	1203	GTTTCTTATAGAAAGGCCATGTACACACTCTGGAAAGCCAGGGTTTCGACTGGCAAG	1265
Oy	1376	CTGTTCCTCCGTGGTCCACACAGCGATGATGATGACACTGCCAAGCCGTGGTGTGAG	1435
Db	1263	ACCTTCTC-----CAGCACTGCTGGAAATCTGTGGAGACCG	1301
Oy	1436	AACAAAGCCATGATTTAATGGGCGCTGTGGGGGCTTCAGGCTTCTGGCCTTAAGGCTTG	1495
Db	1302	CTGAAGGCCATGCTGGAGTGGGCCCCGCGGTGCTTCAACCTCACTGGATGAGGGCTTC	1361
Oy	1496	GAGCCACCAAGAAAGAGAACAATCCCTACAAAGATGTATACACGACATGTGGGTGAA	1555
Db	1362	AAAC-----CAACAAGAAAGCAACCCGAAACAAAAGTGAAGACACACAACTGAC	1415
Oy	1556	CGTAGGACAGCTGCCTACGACACCTTCACGACCCAAAGCCCCGGAAGACAGCAGCG	1615
Db	1416	TCTGTCTCTTCTAGTCCCGCCACCCACCCAGCGCTCCAAACAAATACCTATGGCGGGAAG	1475
Oy	1616	GAGAAGCCCAAGGTCACAAAGATTATTGATGAGCGCAGCAAGAGAGCGGTGTGTACGAG	1675
Db	1476	GACGAGGAGGAGATGAGGAGAGC-----CGAAGACGATGCGTTGTGAA	1520
Oy	1676	GCGGGCAGAAAGTGGCCGAAACATTGAGAGACATGTGCATTCCTGTGGAGACCTCATGTT	1735
Db	1521	GTCAACAAACAAAGGCGCAATGTGGAAGACCGGTGTTTCTGTGGAAAGAAACCT	1580
Oy	1736	ACCCGTGAACCCCCCTGTGTGGTGGAGAAATGTGCAAACTGTGAAGAATGCTTTG	1795
Db	1581	GTGTCTTCCACCCCTCTTTGAGGGTGGGCTGTGTCAAGATGTCCGGATGCTTCTCTA	1640
Oy	1796	GAGTGTGCGTACAGTACAGACGAGAGGGTACCAATCTACTGACATCTGTCTGTGG	1855
Db	1641	GAGCTCTTACATGATGATGATGAGAGCGGTATACATCTCTACTGTACCCTGTCTGTGAG	1700
Oy	1856	GGCCGTGAGGTCTCATGTGCGGAACAACAATCTGTGTGAGGTGCTTTTGTGCGTGGAGT	1915
Db	1701	GCCCGTAACTGCTGTGTGCGAGTGAACAACAGTGTGTGAAGTGTCTGTGTGGAGTGT	1760
Oy	1916	GTGACACTTGTGTGGGGCGGGGGGCTGCCAGGACGCCATTAAAGAGAAGCCCTGGAAC	1975
Db	1761	CTGAGAGTGTGTGTGGCGCCAGCAAGCTAGTGAAGATGCCAAGCTGCAGGAACCTGTGAC	1820
Oy	1976	TGCTACATGTGCGGAGACAAGGGTACTACGGGCTCTCTCGGCGGAGAGGACTGGCC	2035
Db	1821	TGCTATATGTGCTCCCTTAGCGGTGCATGGGGTCTCTCCGACAGAGAAAGATGTGAAC	1880
Oy	2036	TCCCGGCTCAGATGTTCTGTGCTAATTAACAGAC-----AGSAATTTGAACCTCCAAAG	2092
Db	1881	ATGGCGCTGCAAGACTTCTTCACTGATGATCTGACGTGGAAGAAATTTGAGCCACCCAAAG	1940
Oy	2093	GTTTACCCAGCTTCCCAAGCTGAGAAAGGAAGAACCCATCCGGGTGTGTCTCTTGTAT	2152
Db	1941	TTGTAGCCACAAATTTCTGTGACCCAAAGAGAGGCCATTAAGATCTGTCTCTGTTTAT	2000
Oy	2153	GGATTCCTTCAAGGGCTCTGTGTCTGAAGAACTTTGGCATTTCAAGTGGACGCTACAT	2212

Db		2001	GGAATTCGCAAGGGGGTACTTTGGTGCTCAAGAAGTGTGGATTAAAGATGGAAAAAGTCATT	20650
Oy		2213	GCCTCGSAGATGTGTAGAGACATCCATCACCCTGCGGCGCATGTGCGGCCACCAAGGGGAAGATC	22727
Db		2061	GCCTCCCAAAGTCTGTGCACAGTCATCCATCGCTGTGGAAAGCTGTAAAGCATGAAAGGCCAGATC	21250
Oy		2273	ATGTACGTGGGGGACCTCGCACGCGTCAACACAAACATATCCAGAGATGGGGCCCATTC	23322
Db		2121	AAATATGTCAATGACGTCCGGAAAATATCCAACAGAAAATAATTTGAAGAGTGGGGCCCGTTTC	21800
Oy		2333	GATCGGTGTAATTTGGGGGCGAGTCCCTCGCAATGAGACTCTCCATCGTCAACCCCTGTCGCAAG	23952
Db		2181	GACATGGGTATCTGTGTGAAGAGCCCATGCATATGATCTCTCTCAACGTCAATCCTGCGCGCAAA	22450
Oy		2393	GGCCCTACAGAGGGGACACGTGGCGGCTCTTTTGAATTCCTAACCSCCTCCGATATGAGG	24532
Db		2241	GCTTATATATGAGGGCACAGAAAGGCTCTTTTGAGATTTCACACTTGCTGCAATTTATAC	23000
Oy		2453	CGCGCCCAAGAGGAGATATGTCGCCCTTTCTTGCGCTTTTGAGAAATGTGTGGCGCATG	25112
Db		2301	CGCCCCCAAGGGGGCACACACCGTCAATCTCTTGSAATGTGTGAAATGTTGTGGCATG	23600
Oy		2513	GGCGTTAGTACAAGAGGGACATCTCCGCAATTTCTGAGTCCAACCTGTGATGATTGAT	25722
Db		2361	AAAGTAGATGACAAGAAAGACATCTCAAGATTTCTGGCATGTAAACCAAGATGATGAT	24200
Oy		2573	GCCAAAGATGTACAGTCTCACACAGAGGCCCGCTACTTCTTGGGGTAACTTCCCGGTATG	26322
Db		2421	GCCATCAAGGTGTCTCTGCTCTCACAGGGGCCCGGTACTTCTTGGGGTAACTTCCCGGTATG	24800
Oy		2633	AACAGGCGCTGTGGCATCCACTGTGTAATGATMACTGTGAGAGCTGCAGAGAGTCTGTGAGCAT	26922
Db		2481	AACAGGCGCCCTGTATGGCTTCAAAAGAAATGATMACTGTGAGAGACTGCTGTGGAGTTTC	25400
Oy		2693	GGCAGAGATACCCAGTTCAGCAAAAGTAGAGACCATTAACAGAGTCAAACTGCATTAAG	27522
Db		2541	AGTAGAGACACAAAGTTAAAGAAAGTCAACAAATPACACCAAGTGAATCTCATACAA	26000
Oy		2753	CAGGCAAAACACAGATTTCTCTGTCTTGTATAMTAGAAAGAAAGAGACATCTTATGATGTC	28122
Db		2601	CAGGCAAAACACAGTTTTCTCCGTATGATGATMAATGGCAAGAGACAGTTTTGTGTGTC	26600
Oy		2813	ACTGAATGSAAGAGSTATTATTGTTCCTCCAGCTGCATATACTGACGCTCCCAACATGAGC	28722
Db		2661	ACTGAGCTCAAAAGATCTTCGCGCTCTCCCTGCATACAGACATGTCCAAACATGGGC	27200
Oy		2873	CGCTTGGCAGAGGACAGACTCTGTGGCGCGTCAATGSAOCGTGCCAGTCAATCCGCACATC	29322
Db		2721	CGCGGCGCCCGTGAAGAGTGTGTGGAGGCTCTGAGAGCTGACCGGTATACACACTG	27800
Oy		2933	TTCGCTCCGCTGAGAGAGATATTTTGGCTGTGTA	2967
Db		2781	TTTGCCCTTGAAAGACTACTTTGGCTGTGAATA	2815
RESULT 14				
AF151970				
LOCUS		4278 bp	mRNA	linear ROD 08-JUN-2000
DEFINITION		Mus musculus DNA cytosine-specific methyltransferase isoform 2		
(Dnmt3b) mRNA, complete cds.				
AF151970				
VERSION				
KEYWORDS		AF151970.1 GI:8347119		
SOURCE				
ORGANISM		Mus musculus.		
		Mus musculus		
		Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
		1 (bases 1 to 4278)		
REFERENCE		Xlo,B., Chen,Y.T., Zhu,M., Luo,Y.J., Zhu,N., Xu,S.C., Wu,G.Y. and		
AUTHORS		Shen,X.		
TITLE		Cloning of full-length Dnmt3b cDNA and its alternative splicing		
JOURNAL		isoforms in mouse embryonic tissue		
		Unpublished		

REFERENCE 2 (bases 1 to 4278)
 AUTHORS Yin, B., Chen, X. T., Zhu, M., Luo, Y. J., Zhu, N., Xu, S. C., Wu, G. Y. and
 TITLE Direct Submission
 JOURNAL Submitted (17-MAY-1999) Department of Biochemistry and Molecular
 Biology, Institution of Basic Medical Sciences, Chinese Academy of
 Medical Sciences, Dong Dan San Tiao 5, Beijing City 100005,
 P.R.China

FEATURES
 source location/Qualifiers

1..4278
 /organism="Mus musculus"
 /strain="Km"
 /db_xref="taxon:10090"
 /dev_stage="8-9 day old embryo"
 1..4278
 /gene="Dmrt3b"
 412..2931
 /gene="Dmrt3b"
 /function="methylates cytosine in DNA"
 /note="alternatively spliced"
 /codon_start=1
 /product="DNA cytosine-specific methyltransferase isoform 2"
 /protein_id="AA074516.1"
 /db_xref="GI:8347120"
 /translation="MKGDSRLNEEGASGVECTIVNGNFSDDSSDTKXAPSPVLE
 AICTEPVCTPETRGRSSSPASVDMEVEVTPSVSTPVDLSDGDOEGEDTQVASELSY
 LRRETKDTRRSRSPVTRRHSNGTSSLEORASPRITTRKGRHRYHOETVPFPAIR
 SRKRASSSSTPSSPVSVDMEVEVTPSVSTPVDLSDGDOEGEDTQVASELSY
 GDSVEYQDDEKFEIGLWVCKI KFSMPMPAMVSMKATSRQAMPKRWQWQDSEK
 SEISADRLVAGLFSQHNLAENKLVSYRKAMVHTLEKARAGKTFSSPESLED
 OLKPLMEAHAGFKPTGIEGLKPNKQEPKSRRTNDSAESPPPKLKTNSSG
 KDGEDESEREMASVETNKNLIEDRCISGKNRVSPTPEEGGICQSCRPLET
 EYVDDGQSYCTCCGCEGRELICNSCTSCRCFVCELEVLVAGTADEAKLOEPRS
 CYICLQORCHGVLRKKDMRLQDFTDPELEPEPLVPAIPAKRRLRVLSL
 PDIAITVYLAKELIGKVEYLA SEVCAESIAVTVKHEQILYVDRITTKNIE
 WGFEDLVIGSPNDLSNVPARKGLYEGTGLFEYHILANTRPKEGDRPFVWF
 ENYVAMKVNDRKDISRLACNPYIDAIKYSAHRAHYEWNLPGMNRPVMAKNKL
 ELQDLEFSTRLAKIKVORTTTSNSIRGKNLPVYVNGKDYVIMCTELEHIFGFP
 AHTDYSNMGRGRKRLGRMSVPIRHLFAPLKDYFACE"

BASE COUNT 1055 a 1113 c 1154 g 956 t
 ORIGIN

Query Match 18.3%; Score 786.6; DB 10; Length 4278;
 Best Local Similarity 64.5%; Pred. No. 1,3e-150;
 Matches 1261; Conservative 0; Mismatches 649; Indels 45; Gaps 4;

1016 GAAGCCGCTGGGTCATGCTGGGACAGAGATGCCAAGAGCGATGACAGCA 1075
 1018 GATCAGGAGGATATGATACACACAGAGTGATGACAGAGCATATATGAGACAGCA 1077
 1076 GATCAGGAGGAGCGGGGCTTTGGCATTTGGGAGCTGTTGGGGGAACGCGGGGC 1135
 1078 GATATCAGGATATTAAGATTGGAATGATGACCTGCTGGGGAAGATCAAGGCG 1137
 1136 TTCTCTGTGGCCAGCCGCTATGCTGTTGGTGTGATGACGGGCGGAGCCAGCACT 1195
 1138 TTCTCTGTGGCTGCTGATGATGCTGTGTAAGGCGACCTCCCAAGCAGCAGCCATG 1197
 1196 GAAGGACCGCTGGGCTCATGTGCTGAGAGCGCAATTTCTCAGTGTGTGTGAG 1255
 1198 CCGGAGATCGCGGGGTACAGTGTGTTGATGAGCAAGTTTCTGAGATCTCTGCGAC 1257
 1256 AAGCTATGCGCGGTGAGCTGCTTTGTCAGTGGCTTCCACAGGCGCAGTCAACAGCAG 1315
 1258 AAGCTATGCGCTGTGGGCTGTTCACACAGCACTTAATGCTTACTCTTCAATAGCTG 1317
 1316 CCGATATCAGGAGCACTCTGAGAGTCTGTCAGAGTGGCCAGCAGCCGCGGGGAG 1375
 1318 GTTCTTATATGAGAGGAGCATGTACACACTCTGAGAAAGCCAGGGTTCAGCTGGCAG 1377
 1376 CTCTTCCGCTGTGCGACGACGAGATGAGATGACACTGCCAAGGCGGTGAGGTGACG 1435

1378 ACCTTCTC-----CAGCAGTCTGGAGAGTCACTGAGAGCAG 1416
 1436 AACAAGCCCATGATGATGAGGCGGCGCTTCCAGGCTTGTGGCCTTAAGGCGCTG 1495
 1417 CTGAGCCCATGCTGGAGTGGGCGCCACGGTGGCTTAAGCTACTGAGATCGAAGGCTC 1476
 1496 GAGCAGCAGAGAGAGAGAAATCCCTACAAAGATGTATACAGCAGTGTGGTGGAA 1555
 1477 AAGC-----CAACAG 1530
 1556 CCGAGGAG 1615
 1531 TCTGCTCTCTTGTGATGCTCCCCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1590
 1616 GAG 1675
 1591 GACGAG 1635
 1676 GTGCGGAG 1735
 1636 GTACAGCAG 1695
 1736 ACCCTGAG 1795
 1696 GTGCTCTTCCAGCCCTCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1755
 1796 GAGTGTGCTGACAGTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1855
 1756 GAGCTCTTCACTACATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1815
 1856 GGGCGTGAAGTGCATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1915
 1816 GGGCGTGAAGTGCATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1875
 1876 CTGAGAGTGTGAG 1935
 1976 TGCTACTGTGCGGAG 2035
 1936 TGCTACTGTGCGGAG 1995
 2036 TCCGAGCTCAGATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2092
 1996 ATGCGCTGAG 2055
 2093 GTTATCCAGCTGTCCAGCTGAG 2152
 2056 TTGATCCAGCAATTCCTGAG 2115
 2153 GGAATGCTCAG 2212
 2116 GGAATGCTCAG 2175
 2213 GCGCTGAG 2272
 2176 GCGCTGAG 2235
 2273 ATGATGCTGAG 2332
 2236 AATATGCTCAG 2295
 2333 GATCTGCTGATGAG 2392
 2296 GATCTGCTGATGAG 2355
 2393 GGCCTGAG 2452
 2356 GGTATATGAG 2415
 2453 CGGCGCAG 2512
 2416 CGGCGCAG 2475

2513 GGCCTTACGACAGAGACATCTGCGATTTCTCGATCCAGCCTGTGATGATGAT 2572
2476 AAGGATGATGACAAAGAAACATCTCAATATTCCTGGATTAACCCAGATGATGAT 2535
2573 GCGCAAGAGAGTGCATGCGACAGAGGCGCTCTCTCTGCGATGATGATGATGAT 2632
2536 GCGATCAGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2595
2633 AACAGAGCGGTGGATGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 2692
2596 AACAGAGCGGTGGATGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 2655
2693 GCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2752
2656 AGTAGAGACGACAAAGTAAAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 2715
2753 CAGAGCAAGACGACATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2812
2716 CAGAGCAAGACGACATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2775
2813 ACTGAATGAGAAAGGATATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2872
2776 ACTGAGCTGCAAGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2835
2873 CCGTGGCGAGAGCAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2932
2836 CCGGCGGCGCGCGCGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2895
2933 TTGCTGCTGCTGAGAGATATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2967
2896 TTGCTGCTGCTGAGAGATATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2930

RESULT 15
AP156488 4145 bp mRNA linear PRI 05-SEP-1999
LOCUS Homo sapiens DNA cytosine-5 methyltransferase 3 beta 1 (DNMT3b)
DEFINITION
ACCESSION AP156488
VERSION AP156488.1 GI:5823167
KEYWORDS
SOURCE Homo sapiens.
ORGANISM
REFERENCE Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
AUTHORS Xie, S., Wang, Z., Okano, M., Nogami, M., Li, Y., He, W. W., Okumura, K.
TITLE Cloning, expression and chromosome locations of the human DNMT3
JOURNAL
PUBMED
RUIBARD
10433969
2 (bases 1 to 4145)
Xie, S. and Li, E.
Direct Submission
Submitted (03-JUN-1999) CVRC, Mass. Gen. Hospital, 149 13th street,
Charlestown, MA 02129, USA
FEATURES
source
1. 4145
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="20"
/map="between D20S182 and D20S106; 39,9-50.2 cM"
1. 4145
/gene="DNMT3B"
115..2676
/gene="DNMT3B"
/note="de novo DNA methyltransferase"
/codon_start=1
/product="DNA cytosine-5 methyltransferase 3 beta 1"
/protein_id="AAD53063.1"

/db_xref="GI:5823168"
/translation="MKGDTRLHNGEEDAGREDISILVNGCAGSDQSDSPILAEIRTP
EIRGRSSVRLSRKREVSLSLSYTDLDGDDGSDGSGSTPVPMLPILFERTRESEPA
VTRRNNSVSSREHRRPSPRSTRGORNRINIDESPFPATPRLSRATRAAGPWPSP
PSSYLTIDLTDDTDEDTGTSSTPYARLADSOQGMSEPOVEADSGDSDSEYO
DGKEFGIDGLVNGKIKGFSTMPAMVSMKTRKQAMSGMRVOWFGDGKFSVSDAK
LVALGLESQENMLATFNKLVYRRKAMHALEKRVAGKTPSPSGDLEQILKPMLE
MAHGKRPPTGIGLAPNNTOFVYVYKSKVRAGSRLSEKRYENKRRRTADDSATSDY
CPAPRKLKCNVNNNGKDRGDESDQSEDMADVANNNSLSDGLSCGSRKPNVSPHLE
EGELQCTCRDFLELFYMDGQYOSYCTCCBRELILCSITSCRCFECLELVY
GTGTAAEAKLOEPMSCYMLPDRCHGVLRKRMNVRLOAFPTSGLEYEAPRLYLA
IPARRRPPIRLVSLDFGIATGYLVKELGIVKRVYASVESEIATGVTHBSNLY
VNDVRNITRKNIEWNGPDLVIGSPCNDISVNPAPRKVLKYSALHREFTHILNLS
RPRREGDREFEWMFENVAMKRVGRDISHFLSCNPMVTDALKVSAAHRAHFTENGLD
GMRRPVYASKNDKLELDQCLEYENIRAKLRKVOTITTKSKINOGNOLPEVPMGKED
VLNCTELEIRTFSPVHYHDVSNMGAROKLRLRSWSVYIHLPLADYFACE"

BASE COUNT 1041 a 1083 c 1096 g 925 t

ORIGIN

Query Match 17.88; Score 764.4; DB 9; Length 4145;
Best Local Similarity 63.08; Pred. No. 4.5e-146;
Matches 1244; Conservative 0; Mismatches 708; Indels 26; Gaps 3;

1017 AGCCGTGGGGTCCGATGCTGGGACAGAAATGCCCAAGACGCGATGACAGCCAG 1076
1017 AGCAGAGAGACGCGCGGGCTTTGGCATTTGGGAGACGCTGTGGGGAACGTGGGGCT 1136
1077 ACTACAGAGACGCGCGGGCTTTGGCATTTGGGAGACGCTGTGGGGAACGTGGGGCT 1136
761 AGTATCAGAGATGGAAGAGATTGGAAATAGGAGACTGTGTGGGGAAGATCAAGGCT 820
1137 TGTCTGTGGTCCAGAGCGCGCATGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1196
821 TGTCTGTGGTCCAGAGCGCGCATGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 880
1197 AAGGACCGCGTGGTCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1256
881 CCGGAGTGGTGGT 940
1257 AGCTGATGCGCGTGGT 1316
941 AACTGT 1000
1317 CCATGTACGCAAGACCATCTACAGAGTCTGTGACAGTGTGTGTGTGTGTGTGTGTGTGT 1376
1001 TCTCTATCGAAAGCCATGTACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1060
1377 TGTTCGCGGT 1433
1061 CCTTCCCGAGAGCCCTGT 1120
1434 AGAACAGCCCATGATGT 1476
1121 CCCAGGGGGCTTCAAGCCACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1180
1477 -----TTCGTGCGCTTAAGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1531
1181 TGTGTAATAGTCAGAGT 1239
1532 GTGTACAGAGCATGT 1591
1240 GAGAACAGACTGCAAGACCAACAGCTGAGCTGACCTGTGTGTGTGTGTGTGTGTGTGT 1299
1592 AAAAAAGCCCGAAGAGACACAGCGGAGCCCAAGGTCAAGAGATTAATGTGTGTGTGTGT 1651
1300 CCCAAGCGCTCAAGACAAATGTCTTTAACAGGCAAGGCAAGGAGGAGTGAAGATTCAG 1359
1652 ACNAGAGAGCGGT 1711
1360 AGCCGAGAAACAAATGT 1419
1712 ATCTCTGTGGAGCTCAATGTTTACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1771

Db 1420 TTGTCCTGTGCGAGAAAAACCCGTCCTCCACCCTCTTTGAGGGGGGCTCTGT 1479
QY 1772 CAATACTGCAAGACTGCTTCTGAGGTGTGCGTACCAAGACGACGAGGCTACAG 1831
Db 1480 CAAACATGCGGGATGCTTCTGAGCTCTTTTACATGATGATGACGATGCTATCAG 1539
QY 1832 TCCCTACTGCACATCTGCTGTGGGGCCGTGAGGTGCTCATGTGCGAAACAACTGCG 1891
Db 1540 TCTTACTGCACCTGTGCTGCTGAGGGCCGAGAGGTGCTGCTTTCACACAAACAGCTGC 1539
QY 1892 TGCAGGTCTTTTGGCTGAGGTGTGACCTCTTGCTGGGGCCGGGGCTGCCACGCA 1951
Db 1600 TGCCTGTCTGTGTGAGTGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAG 1659
QY 1952 GCCATTAGGAGAACCCCTGGAACCTGCTACATGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAG 2011
Db 1660 GCCAAGCTTCAGAGGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAG 1719
QY 2012 CTGCGGGGCGAGAGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAG 2071
Db 1720 CTGCGGGGCGAGAGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAG 1779
QY 2072 CAGCAATTTGACCTCCAAAGGTTTACCACTGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAG 2131
Db 1780 CTTGAATACGAGCCGCCAAGCTGTACCTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAG 1839
QY 2132 CGGGTGTCTCTCTTGTGATGGAATGCTACAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAG 2191
Db 1840 CGAGTCTGTCTGATGTTGATGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAG 1899
QY 2192 ATTGAGTGGACCTGATGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAG 2251
Db 1900 ATAAAGGTAGGAAGTACGCTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAG 1959
QY 2252 GTCCGGACACAGGGGAAATGATGATGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAG 2311
Db 1960 GTGAGGACAGGGGAAATGATGATGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAG 2019
QY 2312 ATCCAGAGTGGGGCCATGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAG 2371
Db 2020 ATTGAAGATGGGGCCATGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAG 2079
QY 2372 ATGCTAACCTGCTGCAAGGGCTGATGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAG 2431
Db 2080 AATGTGATCCAGCCAGGAAGGCTGATGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAG 2139
QY 2432 TACCGCTCTGATGATGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAG 2491
Db 2140 TACCACTCTGATGATGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAG 2199
QY 2492 TTTGAGATGTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAG 2551
Db 2200 TTTGAGATGTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAG 2259
QY 2552 TCCAACTCTGATGATGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAG 2611
Db 2260 TGTATTCATGATGATGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAG 2319
QY 2612 TGGGGTAACTTCCCGGTATGAAACAGGCTTGGCATCCATGTAATGATTAAGCTGAG 2671
Db 2320 TGGGGCAACCTACCCGGGATGAAACAGGCTTGGCATCCATGTAATGATTAAGCTGAG 2379
QY 2672 CTGAGAGTGTCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAG 2731
Db 2380 CTGAGAGTGTCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAGGTGCTGAG 2439
QY 2732 ACAGGTCAACTTCATTAAGCAAGGCAAGACCATTTTCTGCTTCATGATGAG 2791
Db 2440 ACCAAGTCGAACCTGATCAAAAGGCAAGACCATTTTCTGCTTCATGATGAG 2499
QY 2792 AAAGAGCATCTTATGCTGCTGAAATGAAAGGCTATTTGCTTCCAGTCCACTAT 2851
Db 2500 AAAGAGCATCTTATGCTGCTGAAATGAAAGGCTATTTGCTTCCAGTCCACTAT 2559

QY 2852 ACTGAGCTTCACATGATGAGGCGCTTGGCGAGAGAGACTGCTGGGCGGCTCATGAGC 2911
Db 2560 ACAAGCTGTCCACATGAGGCGGCTGCTGCGCGAGAAAGCTGCTGGAAAGCTCTGGAGC 2619
QY 2912 GTGCCAGTCAACGCGACCTCTTCCCTGCTGAGGAGATTTTGGGTGTGTA 2967
Db 2620 GTGCTGTCAATCCGACACCTCTTGGCGCTGTAAGGACTTATGATGATA 2675

Search completed: July 17, 2003, 07:38:03
Job time : 7201.89 secs

THIS PAGE BLANK (USPTO)

```
1. .3005
   /gene="DNMT3A"
230. .2968
   /gene="DNMT3A"
```

```

/define= DMT3A
/function="de novo DNA cytosine methyltransferase
/codon_start=1
/product="DNA cytosine methyltransferase 3 alpha"
/protein_id="F4ND33084.2"
/db_xref="GI:12746532"

```

/t/nslatla n "W2MPPSSGDDTIVSS3AABEDKSG3L6PDLKLEKRSPOPE
 ARKVGROPRKKHHPPVEGSDTPDPAV3IKSP5MADDSK36L6PDLKLEKRSPOPE
 ESDSPAGGOGGAPPEGECAEATLEP5ASRAVENUCMCPCKRGRAPEALAKREKKA
 MAMESRNLRLGOGJAMESLSLORPMPPLTFOAGDPVJ5KRRKAPALAKREKKA
 KXIAAMNVAENQOPGQSSQLESP5PAVOPDPDAPSPVATTPPEVSD3GAKNARK
 AGDDEPEXEDRGIGLEIWLXRGCSMPG3L15MMTGV3ASAGL6PVDHSDSE
 KSV5CWEKLPAL5FCSA5HQAQYNQMPRALEYV5GVAASAGL6PVDHSDSE
 ST3AAVEYQNPATL5MNLG6FOP3GKLEPPEEKNNIKEYTDNVAPEALAYV
 PPAPKPRKRAEATLEI5DETRERATLEYV5GVAASAGL6PVDHSDSE
 PVGKCMCKNCEKLECAVOYDD25O5CT1CGG6VEALCMMNNCKRCFEV5CYDLL
 VPGQAQA4KED6MNYCMGHCKT5L6LRREDDWSYLRJ6PFAHHQK6PVPVTH
 PVAPEAKRIRYLSL7PDLATGDLV5LWDL6LQVDRY1ASVED2ITGVANRHOGRK
 YGVDV5YQKID6WGFVLDVIGGSEFNDLSVYNARGLA5EGSRLEPFEYLRHLH
 ARKEDDEPFWMLEENVA5GSDKDRID5KFL5E5NP5IDA5K5AHRAV5PYN5N
 PCMNPR1A5VNDLELOE6GSDKDRID5KFL5E5NP5IDA5K5AHRAV5PYN5N
 DILWCTEMR5YCGPVRHTDVS5N5RLAROLLR6G5SV5VIRHL5PAPLEK5YACV"

Query Match	70.0%;	Score 3005;	DB 9;	Length 3005;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3005;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0

1 GCCGCGGCAACAGGGCGGCGAGCCGCGGCGCCGACCCCAACGGCGGTGAGCC 60

Db

1 GCCCGGGCACCCAGGGCGGCAGCCGGGCGCCCGACCCACCCGGCATACGGTGAGCC 60

61 ATCGAAGCCCCCACCACAGAGGCTGACAGAGGACCCTTCACCAGAGGGCTCACACCCGG 120

61 ATCGAAGCCCCCACCACAGGCTGACACAGGCACCGTTCACACAGAGGGCTCACACACGGG 120

121 ATCTAAGTTTAACTCTGCGCTCCAAGACACAGATAATTCCCTCCCAAGCC 180

[illegible]

b5 121 A C C A G A T C C C G G C C A G A T C C C G G C C A T 240

[illegible]

Db
181 CAGCAGCCCCCCAGCCCCGCGACGCCCAACGACGGGAAA 306

[illegible]

241 GCCTCCAGCGGCGCGGAGACCAAGACACTCTGCTGCGAGGCGGAGGAGGACCGAAA

301 GGACGGAGAGAGCAGGAGACCCGCTGGCAAGGAGGAGCGCCAGAGGCCACACACAC 302

301 GGACGGAGAGCAGGAGAGCCCGTGGCAAGGAGGCGCCCAAGAGCCCCAGCAC

361 GGCACGGAAGTGGGGGGGCTTGGGAGGAAAGCCAGACCCCGGTGGAAGCGGTGA 42

Db 361 GGCACGGAGGTGGGGCGGCTGTGGAGGAGCGCAGCACCACCCCGGTGGAAAGCGGTGA 42

421 CACGCCAAGGACCCCTCGGTGATCTCCAGTCCCATTCATGGCCAGGACTCAGGCGC 48

421 CATTGCTAAAGGAGCCCTGCGGTGATCTCCAGTCCCATTCATGGCCAGGACTCAGGGC 48

481 CACACACCTATTATACCAATGGGACCTTGGAGAACCGAGTGAACCCCAACCCAGAGAGGG 54

[illegible]

D6 481 CTCAGAGCCTAATACCCCAATGCGGCATGGTTCAGACTGAGAC 60

541 GAGCCCTGGCTGGGCGCAAGGAGAGAGGCTGAGCTGAGAC 60

(Db) 541 GAGCCCTGCTGGGGGGCAGAGGCGGCCCCCAACACACGCCCGAAG 66

601 CCTGCTGAAGCCTCAAGAGCAGTGGAAAAATCGGTTCCTCCACCCCCCAAGGAGGCCAAC

```
1. .3005
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /chromosome="2"
   /map="2p23: between D2S171 and D2S174; 48-50 cM"
```

THIS PAGE BLANK (USPTO)

Db 601 CTGCTGAAGCCTCAAGAGCAGTGAAAAATGGCTGTGTCACCCCAAGAGAGGCGGAG 660
 Oy 661 AGCCCTCTGAGAGAGGGGCAAAAGAACAGAGAGACCAATCATGAAATCCATGAAATGGA 720
 Db 661 AGCCCTCTGAGAGAGGGGCAAAAGAACAGAGAGACCAATCATGAAATCCATGAAATGGA 720
 Oy 721 GGGCTCCCGGGGCGGCTGGGGGGTGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGG 780
 Db 721 GGGCTCCCGGGGCGGCTGGGGGGTGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGG 780
 Oy 781 CATGCCGAGGCTCACCTTCCAGGCGGGGACCCCTACTACATCAGCAAGCCGAGCGGGA 840
 Db 781 CATGCCGAGGCTCACCTTCCAGGCGGGGACCCCTACTACATCAGCAAGCCGAGCGGGA 840
 Oy 841 CGAGTGGCTGGCAGCCTGGAAAAAGGAGGCTGAGAAAGCAAGGCTATTGCAAGAAAT 900
 Db 841 CGAGTGGCTGGCAGCCTGGAAAAAGGAGGCTGAGAAAGCAAGGCTATTGCAAGAAAT 900
 Oy 901 GAATGCTGGAGAAAAACAGGGGCGGGGGAGTCTGAGAGTGGAGGAGCGGAGCC 960
 Db 901 GAATGCTGGAGAAAAACAGGGGCGGGGGAGTCTGAGAGTGGAGGAGCGGAGCC 960
 Oy 961 TCTGCTGTGACAGAGCCCACTGACCCCGCATCCCACTGTGGCTTACCAAGCCTGAGCC 1020
 Db 961 TCTGCTGTGACAGAGCCCACTGACCCCGCATCCCACTGTGGCTTACCAAGCCTGAGCC 1020
 Oy 1021 CGTGGGAGTCCATGCTGGGGGACAGAAATGCCAAGAGAGGCTGACAGAGCCAGAGTA 1080
 Db 1021 CGTGGGAGTCCATGCTGGGGGACAGAAATGCCAAGAGAGGCTGACAGAGCCAGAGTA 1080
 Oy 1081 CGAGAGCGGCGGGGCTTTGGCATTTGGGAGCTGGTGGGGGAAACAGCGGGGCTTCTC 1140
 Db 1081 CGAGAGCGGCGGGGCTTTGGCATTTGGGAGCTGGTGGGGGAAACAGCGGGGCTTCTC 1140
 Oy 1141 CTGGTGGCAGGCGGCATTTGTCTTGTGTGATGACGGGCGGAGCCGAGCAGCTGAAAG 1200
 Db 1141 CTGGTGGCAGGCGGCATTTGTCTTGTGTGATGACGGGCGGAGCCGAGCAGCTGAAAG 1200
 Oy 1201 CACCGCGTGGCTCATGTGTTGGAGAGAGGCAAAATTCATAGTGTGTGTGAAGAGCT 1260
 Db 1201 CACCGCGTGGCTCATGTGTTGGAGAGAGGCAAAATTCATAGTGTGTGTGAAGAGCT 1260
 Oy 1261 GATGCGCGTGAAGCTGTTTGTGAGTGGCTTCCAGCAGGCGCAGTAAACAGAGCCCAT 1320
 Db 1261 GATGCGCGTGAAGCTGTTTGTGAGTGGCTTCCAGCAGGCGCAGTAAACAGAGCCCAT 1320
 Oy 1321 GTACGCAAAAGCCATCTACAGAGTCTCAGAGTGGCCAGCAGCCGCGGGGAGAGCTGTT 1380
 Db 1321 GTACGCAAAAGCCATCTACAGAGTCTCAGAGTGGCCAGCAGCCGCGGGGAGAGCTGTT 1380
 Oy 1381 CCGCGTGGCCAGCAGCAGTGAAGAGTGAACCTGCCAAGGCCGTGGAGGTGCACAAACA 1440
 Db 1381 CCGCGTGGCCAGCAGCAGTGAAGAGTGAACCTGCCAAGGCCGTGGAGGTGCACAAACA 1440
 Oy 1441 GCCCATGATGAAATGGGCTTGGGGGCTTCCAGCCTTCTGGCCTTAAGAGGCGCTGAGGC 1500
 Db 1441 GCCCATGATGAAATGGGCTTGGGGGCTTCCAGCCTTCTGGCCTTAAGAGGCGCTGAGGC 1500
 Oy 1501 ACCAGAGAGAGAAATCCCTACAAAGAGTGAACAGCAGATGGGGTGAACCTGA 1560
 Db 1501 ACCAGAGAGAGAAATCCCTACAAAGAGTGAACAGCAGATGGGGTGAACCTGA 1560
 Oy 1561 GGCAGCTGCTACGACACACCTCCACAGCCAAAAAGCCCGGAGAGAGCAGCGGAGGA 1620
 Db 1561 GGCAGCTGCTACGACACACCTCCACAGCCAAAAAGCCCGGAGAGAGCAGCGGAGGA 1620
 Oy 1621 GCCCAAGGTCAAGAGATTTATGATGAGCGCACAGAGAGCGCTGTGTACAGAGTGGC 1680
 Db 1621 GCCCAAGGTCAAGAGATTTATGATGAGCGCACAGAGAGCGCGCTGTGTACAGAGTGGC 1680
 Oy 1681 GCAGAGTGGCGGAACATTGAGAGCATCTGATCTCTGTGGAGCCTCAATGTTACCT 1740
 Db 1681 GCAGAGTGGCGGAACATTGAGAGCATCTGATCTCTGTGGAGCCTCAATGTTACCT 1740

Db 1681 GCAGAGTGGCGGAACATTGAGAGCATCTGATCTCTGTGGAGCCTCAATGTTACCT 1740
 Oy 1741 GGAAACACCCCTCTTCTGTGGAGAAATGTGCCAAACTGCAAGAACTGTTTCTGAGATG 1800
 Db 1741 GGAAACACCCCTCTTCTGTGGAGAAATGTGCCAAACTGCAAGAACTGTTTCTGAGATG 1800
 Oy 1801 TGGCTACAGTACGACGACGAGGCTTACAGCTTACTGACACATCTGCTGGGGGGCGG 1860
 Db 1801 TGGCTACAGTACGACGACGAGGCTTACAGCTTACTGACACATCTGCTGGGGGGCGG 1860
 Oy 1861 TGAGGTGCTCATGTGCGGAAACAAACAACTGCTGCAAGTGTCTTTGCTGAGATGTGGA 1920
 Db 1861 TGAGGTGCTCATGTGCGGAAACAAACAACTGCTGCAAGTGTCTTTGCTGAGATGTGGA 1920
 Oy 1921 CCTCTTGGTGGGGCGGGGCTGCGGAGCAGCATTAAAGAAAGACCTTGAACCTGCTA 1980
 Db 1921 CCTCTTGGTGGGGCGGGGCTGCGGAGCAGCATTAAAGAAAGACCTTGAACCTGCTA 1980
 Oy 1981 CATGTGGGGGCAAGAGGTACCTACGGGCTGCTGGCGGCGGAGAGAGACTGGCCCTCGG 2040
 Db 1981 CATGTGGGGGCAAGAGGTACCTACGGGCTGCTGGCGGCGGAGAGAGACTGGCCCTCGG 2040
 Oy 2041 GCTCCAGATGTTCTTCTGCTTAATTAACAGACAGCAAGATTTGACCCCTCCAAAGTTTACC 2100
 Db 2041 GCTCCAGATGTTCTTCTGCTTAATTAACAGACAGCAAGATTTGACCCCTCCAAAGTTTACC 2100
 Oy 2101 ACCTGCTCCAGTGAAGAGAGAACCCATCCGGGTGCTGTCTTCTTGTGATGAAATGCG 2160
 Db 2101 ACCTGCTCCAGTGAAGAGAGAACCCATCCGGGTGCTGTCTTCTTGTGATGAAATGCG 2160
 Oy 2161 TACAGGGCTCTGTGGTGTGAAGAGTGGGCTTGAAGTGAAGGACCCCTACATGCTCGGA 2220
 Db 2161 TACAGGGCTCTGTGGTGTGAAGAGTGGGCTTGAAGTGAAGGACCCCTACATGCTCGGA 2220
 Oy 2221 GGTGTGTGAGACTCCATCAGCGTGGGAGATGTGGCGGACCAAGGGGAAAGATCATGTAGCT 2280
 Db 2221 GGTGTGTGAGACTCCATCAGCGTGGGAGATGTGGCGGACCAAGGGGAAAGATCATGTAGCT 2280
 Oy 2281 CGGGAGACGCGGACGCTCAGACAGAGATTTCCAGAGAGTGGGGCCCTTGTGATCTGCT 2340
 Db 2281 CGGGAGACGCGGACGCTCAGACAGAGATTTCCAGAGAGTGGGGCCCTTGTGATCTGCT 2340
 Oy 2341 GATTTGGGGCACTCCCTGCAATGACCTGCTCAATGCAACCTGCTGCGAAGGCGCTCA 2400
 Db 2341 GATTTGGGGCACTCCCTGCAATGACCTGCTCAATGCAACCTGCTGCGAAGGCGCTCA 2400
 Oy 2401 CGAGGCACTGGCGGCTCTTCTTGAAGTTTACCGCTCTGATGATGCGGCGCA 2460
 Db 2401 CGAGGCACTGGCGGCTCTTCTTGAAGTTTACCGCTCTGATGATGCGGCGCA 2460
 Oy 2461 GGAGGAGATGATGCGCCCTTCTGCTGCTTGAAGTGTGGGCGCATGGGGGTTAG 2520
 Db 2461 GGAGGAGATGATGCGCCCTTCTGCTGCTTGAAGTGTGGGCGCATGGGGGTTAG 2520
 Oy 2521 TGACAAAGAGGAGCATCTGCGATTTCTGAGTGCACAAACCTGTGATGATGATGCCAAAGA 2580
 Db 2521 TGACAAAGAGGAGCATCTGCGATTTCTGAGTGCACAAACCTGTGATGATGATGCCAAAGA 2580
 Oy 2581 AGTGTACGTGACACAGAGGCGGCTTACTTCTGGGGTAACTTCCCGATGAAAGAGCC 2640
 Db 2581 AGTGTACGTGACACAGAGGCGGCTTACTTCTGGGGTAACTTCCCGATGAAAGAGCC 2640
 Oy 2641 GTTGCATCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2700
 Db 2641 GTTGCATCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2700
 Oy 2701 AGCCAAAGTTTCAAGAAAGTGAAGACATTAACAGAGTCAAACTCCATTAAGCAGGGCA 2760
 Db 2701 AGCCAAAGTTTCAAGAAAGTGAAGACATTAACAGAGTCAAACTCCATTAAGCAGGGCA 2760
 Oy 2761 AGACCAAGCATTTTCTCTCTTCAATGAATGAAGAGACATTTATGATGATGATGATGAT 2820
 Db 2761 AGACCAAGCATTTTCTCTCTTCAATGAATGAAGAGACATTTATGATGATGATGATGATGAT 2820

THIS PAGE BLANK (USPTO)

Tue Jul 22 16:21:56 2003

us-09-720-C

QY 2821 GGAAGGGTATTGTTGTTCCAGTCCACTTACTGAGTCTCCACATGAGCGCTTGGC 2880
Db 2821 GGAAGGGTATTGTTGTTCCAGTCCACTTACTGAGTCTCCACATGAGCGCTTGGC 2880
QY 2881 GAGCGAGAGACTGCTGGCCCGTGCATGAGCGTGCAGTCAATCCGCCACTCTTGGCTCC 2940
Db 2881 GAGCGAGAGACTGCTGGCCCGTGCATGAGCGTGCAGTCAATCCGCCACTCTTGGCTCC 2940
QY 2941 GCTGAAGAGTATTGCTGCTGTGTAAAGGACATGGGGCAACTGAGTAGCGACACA 3000
Db 2941 GCTGAAGAGTATTGCTGCTGTGTAAAGGACATGGGGCAACTGAGTAGCGACACA 3000
QY 3001 AAGTT 3005
Db 3001 AAGTT 3005

RESULT 3
AF068625

THIS PAGE BLANK (UBPTO)